

3 June 1982

Squabbles over East-West pipeline

The Soviet plan to build a huge pipeline for natural gas for customers in Western Europe has irked the United States Administration. But Washington's fears are exaggerated.

The Western economic summit that opens tomorrow (4 June) at Versailles is likely to be more than usually contentious. President Ronald Reagan, one of the several first-time participants at these occasions, is unlikely to find the meeting congenial, if only because the United States will find itself hemmed firmly into a corner about its distinctive (some say eccentric) economic policy, a curious blend of Keynes and Friedman whose chief external effect is to keep international interest rates so high that economic recovery will be impeded. (The Prime Minister of Japan, Mr Zenko Suzuki, will, on the other hand, be pilloried for what he may fairly claim is a paradoxical reason — interest rates notwithstanding, Japan has been too successful.) But when the going gets rough, the United States will reply with its now long-standing complaint that Western Europe should not be preparing to buy natural gas from the Soviet Union. This issue, about which Washington has been sucking its teeth off and on for the best part of a year, is without doubt important. But not in the sense in which the United States appears to believe.

The issue exists because of the discovery of substantial reserves of natural gas in the Soviet Urengoi field, in the southern Urals. The plan is to ship large quantities of this gas westwards towards Europe (East and West) by means of a huge pipeline on which construction is about to start. West Germany, France and Switzerland have already signed contracts to purchase natural gas over the next thirty years or so. American objections to European involvement in this scheme take three forms. First, the argument goes, Western European dependence on natural gas from the Soviet Union will undermine its strategic independence. Second, the method by which the pipeline will be financed, with the eventual customers providing credits for the construction, being repaid with gas at a later stage, makes it possible for the Soviet Union to avoid what would otherwise be a nasty brush with the international capital markets. And, third, some of the heavy pipe-laying machinery to be used in the construction of the pipeline will provide the Soviet Union with access to Western technology in a fashion that will ultimately rebound to its strategic and military advantage.

On all three counts, the fears, while not groundless, are exaggerated. The most obvious danger is that Western European states which come to depend on Soviet gas supplies will then be uncomfortably in Soviet hands if, at some point in the future, the Soviet Union should decide arbitrarily to turn off the taps. The Soviet Union has, however, an enviable record for acting meticulously in the performance of commercial contracts, so that fears of strategic vulnerability presumably attach to the hypothesis that there might one day be a European war in which the Soviet Union was at loggerheads with its customers. Then, no doubt, the gas would be cut off but that would not bring any of the intending customers to its knees (although the war might). In no case is the amount of gas being purchased more than five per cent of the customer's use of energy, while the arrangements that prudent gas consumers will make to insure against a mechanical failure of the pipeline would no doubt also allow them to get by if hostilities should break out. Thus the chief strategic significance of the pipeline is that it would shorten the period of time during which Western Europe could fight a conventional war in Europe, and thus quicken the trigger of nuclear escalation. But this is

entirely in conformity with the military strategy of Western Europe, not entirely welcome to the Soviet Union.

What the United States objections to the pipeline project overlook is the value to Western Europe of a supply of fuel which, although not dominant in the energy strategy of any potential customer, is at once relatively cheap and also secure. Not often, these days, can people hope to get their hands on a supply of energy over a period of three decades. And even the method by which the project is being financed, requiring that the customers should somehow help with the cost of construction, is not the simple gift it is made out to be. For who among the West German steelmakers and their workers, who are supplying pipes, or in the French companies from whom turbines have been ordered, will complain at the extra work that is falling in their laps? But is it not against the interest of the West that European states should make deals with the Soviet Union whose effects are to make the Soviet Union economically stronger? This is the unspoken question embedded in the American objection. The difficulty, of course, is that trade deals willingly arrived at do invariably benefit all partners. One of the continuing and deep-seated differences between the United States Administration and many governments of Western Europe turns on the European view that an economically healthy Soviet Union is a better neighbour than one perpetually on its knees.

The transfer of technology is in much the same case. Knowledge of how huge pipe-laying machinery works would no doubt be an advantage to the Soviet Union in the development of some of its remote fields of oil and gas, but that does not by itself imply that the world becomes a more dangerous place. The more serious of the United States' objections to what is proposed, however, is the decision that some of the high-capacity compressor turbines to be installed in the pipeline embody technology that could be of military value. This appears to be the reason why General Electric, the United States manufacturer, has asked its French licensee not to accept orders for the pipeline project. The decision is unlikely to make sense, for the same equipment can readily be had from elsewhere, for which reason it is hard to think it can be of crucial military value. But in any case, if President Reagan should give the pipeline project an airing in Versailles, he is likely to find M. François Mitterrand, the President of France, taking a sharp but understandable line about the turbines — and the underlying principle.

Where now for UNEP?

On its tenth birthday the United Nations programme for the environment is stumbling.

The United Nations Environment Programme (UNEP), the principal offspring of the environmental jamboree at Stockholm in 1972 and a child of the 1960s, is in another bout of trouble. This week, the governing council of UNEP concluded its annual meeting in Nairobi, where it considered — among other things — the budget for the next two years. Predictably, there was a lot of heat and little light, but at least no delegation threatened to withdraw.

The budget is a constant worry, principally because of wavering by the United States, which has contributed \$7 million a year since



1973, more than a third of the total programme cost. Last year's fear that the United States would contribute nothing at all did not materialize. The State Department offer of \$2.2 million was increased by Congress to \$7.85 million — but UNEP is still waiting for the cheque. This year, the Administration's offer is \$3 million, but Congress may again vote more. Whatever happens, and in spite of the surprise decision of the British government to increase its contribution by a quarter, UNEP's hope of doubling its budget (\$17 million in 1981) now seems a dream. So too must seem the first five years, when UNEP raised \$60 million out of its target of \$100 million within just a few months. The peak budget (\$21 million) was in 1977, since when, in spite of solid support from Scandinavia and the Netherlands, UNEP has gone downhill in strictly financial terms.

In part, this is a negative consequence of the developing countries' growing attention to UNEP as a source of development aid. One of UNEP's more successful exercises, the regional seas programme, began with Mediterranean pollution (which stems mostly from the French Rhône, Italian Po and Spanish Ebro rivers) but has now shifted to more southern seas. This shift of attention is not just a UNEP quirk. The environmental impact of poverty looms over that of affluence. Growth is in fact faster in some countries of the developing world than in the 'North', and where industrial development takes place, it is often without controls. The air of Calcutta is now as bad as that of London in the smogs of the early 1950s, while London has clean air.

Yet the Stockholm conference did not plan UNEP's \$20 million a year to be an environmental action fund for the UN. Rather, Stockholm saw UNEP as a coordinating agency within the UN system, drawing together and rationalizing the actions of the other UN agencies. But the UN agencies — as always — proved jealous of their powers. Moreover, the governing council of UNEP, which represents subscribing nations, ended up dealing almost exclusively with UNEP's small direct programme and budget, ignoring the wider role recommended by Stockholm.

In the event, UNEP's programme comes to rest mainly on the Global Environment Monitoring System (GEMS); Infoterra (UNEP's information system); the International Register of Potentially Toxic Chemicals (IRPTC); and the Regional Seas Programme.

The two last are generally thought to be the most successful. They are based in Geneva, which gives them better relations with the other UN agencies (through the strong UN presence there), and they also happen to be headed by active and committed people: Dutch biologist Jan Huisman for IRPTC, and Yugoslav marine biologist Stjepan Keckes for Regional Seas. IRPTC is on the point of producing its first batch of complete data profiles of 330 chemicals of environmental significance — but it has been severely hampered by a halving of its budget. (It relies now on Huisman, an assistant, and a secretary.) Regional Seas is a little bigger (it has a professional staff of six), has fared better for funding with UNEP, and has scored a number of remarkable successes in the form of treaties controlling the pollution of coastal waters, notably the 'Athens protocol' on land-based sources affecting the Mediterranean. One reason for this success has been Keckes' decision to rely on local scientists — who often must be trained first — to assess pollution in an area, on the principle that governments are more willing to act on advice from their own scientists than on that of a fly-by-night international team. The programme is now working on eight other regional seas around the world.

Of the other programmes, GEMS — the global monitoring exercise recommended by Stockholm — has suffered from lack of resources. A 'colossal effort' is needed to strengthen this, says UNEP in a recent publication. But an independent expert review of the 1972–82 decade, commissioned by UNEP, says that 'even an imperfect series of environmental statistics would be an improvement on present ignorance'*. And Infoterra, the information system, is barely used.

UNEP's work on deserts — another great Stockholm concern — has also achieved little, not so much because of UNEP as because industrialized states have not wished to pay. The UN Conference on Desertification held in Nairobi in 1977 produced good scientific papers and a plan of action; but the UN special account to combat desertification, set up after the conference, has remained almost empty. By the end of 1981 total contributions stood at \$5,000, all from a single contributor, Mexico.

This is not the sum of UNEP efforts, however. It can cite work on soils and pest control; its role in the World Conservation Strategy, which is seen as one of the successes of 1970s environmental management; reports on the impact of energy technologies, and a conference on renewables; its task force on pollution and human health (together with the World Health Organization); its promotion of a series of expert meetings on the preservation of tropical forests; international monitoring schemes for acid rain, and pollutants in food and fresh water. The snag is that this part of UNEP's work is too much of a miscellany to make a mark.

What next for UNEP? Apart from the cash problem, UNEP is concerned with defining its programme for the rest of the decade. For the first time, the major UN agencies have got together to define a joint policy: the 'System-wide Medium-term Environment Programme' (SWMTEP) which covers the period 1984–89. But UNEP played little part in developing it, and it has been described as bland and highly generalized. However, it distinguishes the roles that the different agencies should play, even if it is short on recommendations for specific action. And it may help UNEP find its way through the UN thicket rather better than it did in the 1970s or it will, if this week's governing council gives UNEP the necessary backing.

Universities in limbo

The British government's wish to save money on universities is forgivable; its way of doing so is not.

British universities now know that there will be no reprieve from the fate decreed for them by a thoughtless government. The announcement last week by the University Grants Committee (*Nature* 27 May, p.258) that university budgets for 1982–83 will be essentially those forecast a year ago will have dispelled hopes that the government might not be serious. The result is that contraction must continue for another two academic years. There are two views of what the system will then be like. The government seems ideologically committed to the belief that leaner means fitter. Universities will have been forced to cut what is called 'fat', and some will also have been forced to keep essential activities alive with help from local industry that will reduce demands on the public purse and be worthwhile in themselves. What can be wrong with that, the government asks?

Plenty, is the simple answer. The second and more gloomy view of the future starts from the observation that the contraction of the university system has been ordained within arbitrary constraints. Last year's distribution of grants and, more important, the assignment of targets for student populations, has compelled some universities to shrink more rapidly than others. And the restrictions on student numbers mean that no university can hope to work its way out of trouble by being more efficient — taking more students. Yet demand for higher education in Britain continues to increase, with the result that students are shunted arbitrarily towards other institutions, the polytechnics for example. All this arbitrariness arises, however, not from a belief that restricting access to higher education is in itself virtuous but because the government has not had the wit to devise a scheme for limiting the cost of maintenance grants for students, which are paid by local authorities and charged back to the central government.

To seek to economize is within the government's gift, and may even be its duty. To attempt to do so without solving a simple book-keeping problem, and while barriers persist between the two halves of higher education, amounts to maladministration.

* See *The World Environment 1972–82*, edited by Martin Holdgate and others (Tycoby International, Dublin, 1982).

Clouds may lift over Brookhaven

Congress moves to reprieve ISABELLE plan

Washington

ISABELLE, the new particle accelerator at Brookhaven National Laboratory on Long Island, may be saved for at least another year. Recently, serious problems arose with the original magnets, necessitating major changes in design, and causing delays and a rise in the project's construction cost from \$275 million to \$500 million. And because the troubled ISABELLE began taking a bigger than expected bite out of the US high-energy physics budget the question of whether to finish it at all has become a burning issue among US physicists.

It is also a burning issue in Congress, which is sensitive to pleas from the Long Island community to rescue the big machine, and, by extension, the Brookhaven Laboratory itself. The President's 1983 budget request included no funds for the construction of ISABELLE, although approximately \$24 million was included for work on alternative magnets and other research and development projects. Moreover, a high level panel representing the national physics community had just said that ISABELLE should not be finished unless the overall budget was increased. So ISABELLE seemed doomed.

Or so it seemed until 7 May, when Congressman Tom Bevill (Democrat, Alabama), chairman of a key House appropriations subcommittee which has jurisdiction over the high-energy physics budget, toured the buildings at Brookhaven and declared himself "impressed". Better to see the creature, he said, than read about her in reports. Bevill assured Congressman William Carney (Democrat, New York) that he would try to get as much as \$10 million in construction funds added to the budget when it comes before his subcommittee. Both Carney, who represents the Brookhaven area, and Thomas Downey (Democrat, New York) want to see ISABELLE finished almost as much as the scientists at Brookhaven do.

But ISABELLE has received only a qualified endorsement from other quarters. Congressman Don Fuqua, chairman of the House Science and Technology Committee which advises on how high-energy physics money should be distributed, has recommended either proceeding with ISABELLE or building a less costly machine at Brookhaven and using the ring tunnel already built.

Fuqua's recommendation carries some weight with Congress, as his committee

knows as much about science as any on Capitol Hill. Although it may be construed as an endorsement for ISABELLE, in reality it steers carefully among the factions that have emerged on the question.

More guarded still was the report of a panel headed by George Trilling of the University of California at Berkeley, appointed to evaluate future plans for high-energy physics. Possibilities considered included proceeding with ISABELLE, cancelling ISABELLE to keep other machines going at other laboratories, and building something else. One of the options was a lower cost machine at Brookhaven, and in January the Trilling panel endorsed this recommendation when it reported to the High Energy Physics Advisory Panel of the Department of Energy. Thus, the Trilling group's view carries the weight of the community at large.

The Trilling report said that ISABELLE should be completed only if funds for high-energy physics in the future "averaged" better than the 1982 level of \$395 million, for then ISABELLE's increase would not

hurt other projects. However, since the President has requested \$429 million for 1983, which is equal to the \$395 million level in 1982 dollars, then, according to the advisory panel reasoning, ISABELLE should not be built.

All concerned are trying to avoid an outright showdown between the major US accelerator communities involved in the scramble for funding commitments for future big machines. The issue, they all say, is US leadership in high-energy physics in the 1990s.

The argument is partly about what kind of physics should be done, partly about how experiments are shaped by the machines themselves and their costs and timetables, and partly about keeping the eastern portion of the country, in which Brookhaven is the only remaining big accelerator facility, on a par with the facilities in the Mid-West and the West. Congressman Bevill, it seems, is proposing to resolve all this with a stroke of his pen when the item comes before his subcommittee in the next month or so.

Deborah Shapley

Ulcer cure sweetener for Searle

By the end of this year G. D. Searle and Co. Ltd expect to be testing their first recombinant DNA product in humans. It will be a peptide synthesized according to the instructions carried by a synthetic gene placed inside bacteria. Searle and Imperial Chemical Industries Ltd, initiators of the joint project, hope that the peptide, which is closely related to a gastrointestinal hormone, will prove to be an effective inhibitor of gastric acid secretion and hence of ulcers.

That is one indication that Searle's early attempts to capitalize on biotechnology have survived two setbacks. The first was the departure, in 1980, of two key members of staff to help start up Celltech. Searle's second setback was its failure to make a commercial success of producing beta-interferon from cultured fibroblasts.

Another indication of Searle's resilience in biotechnology came last week with the opening of a new £7 million biotechnology pilot plant at their High Wycombe site in the United Kingdom. The plant will come into operation next month, less than two years after construction started, despite a change in plan half way through. The change involved increasing the part of the plant designed for bacterial culture at the expense of that for cell culture. That shift in balance reflected both the failure of the beta-interferon project and, according to Brian Richards, vice-president of UK pre-clinical research and development, the discovery of several peptides of potential value that could best be produced by genetically engineered microorganisms.

The pilot plant still has facilities for the culture of cells but Searle anticipates more activity in the bacterial fermenters, the largest of which has a 450 litre capacity. The plant, however, is designed to accommodate a 4,000-litre fermenter, large enough to advance from pilot to manufacturing scale in some circumstances. For yet larger scale fermentation, Searle will

ULCER RESEARCH DEPARTMENT



"That's right; we both create and cure them"

have either to build new plant or to arrange to use the vast fermentation facilities of the Japanese company Meiji Seika, using a recent agreement to exchange technology.

That, however, is for the future, when and if Searle has something worth fermenting. Bacterial products planned, apart from the anti-ulcer peptide, include interferons. Much effort is being directed towards making structural variants of interferons with greater or more specific activity than the natural compounds. If

any are discovered the plan is to manufacture them with bacteria containing synthetic genes.

Searle may also use biotechnology in the less glamorous pursuit of producing, more cheaply than at present, the two amino acids that are fused together to make aspartame, Searle's low calorie sweetener which was approved for tabletop use in the United States last year. Both amino acids come from bacterial fermentation and so it is possible that the bacteria could be genetically manipulated to produce either amino acid more efficiently than at present.

Peter Newmark

Remote sensing

What war use?

Landsat, the US remote sensing satellite launched by the National Aeronautics and Space Administration (NASA) could have been used by the Argentine and British governments to obtain images of the Falkland Islands, but probably to little effect. While Landsat can pick up data from all over the globe, it passes over the Falklands only every eighteen days and a further twenty-four hours are then required to process images. Landsat has a resolution of 80 metres and could pick out — with luck and favourable weather conditions — an object as large as an aircraft carrier.

Landsat-3 is a near-polar orbiting spacecraft launched in March 1978. It carries visible and infrared wavelength sensors which send back images used in agricultural surveys, mineral and oil exploration and pollution monitoring. NASA collects the satellite data at Goddard Space Flight Center in Greenbelt, Maryland and it is then sold by the Interior Department through the Geological Survey's EROS (Earth Resources Observation Satellite) Data Center in Sioux Falls, South Dakota.

Fresh images can be obtained from NASA by prior arrangement but the usual time-lag is anything from one to six months. Ground stations within a 2,000 km range can pick up real time broadcasts by radio link. The British defence department's National Remote Sensing Centre at Farnborough, Hampshire, is licensed under ESA's Earthnet programme to tune into transmissions over Europe. Argentina also has a ground station at Mar Chiquita (600 km north-west of Buenos Aires), but to obtain fresh images of the Falklands, it would have to make a special arrangement with NASA so that the data would not be wiped off the recorders on board, which are limited in the amount of information they can store.

The delay in obtaining fresh images and the coarseness of their resolution makes the strategic use of Landsat dubious. Landsat images may, however, have provided information about prevailing terrain conditions.

Jane Wynn

Nairobi environmental meeting

More and less

A Japanese environmental delegation is this week trying to recover its pride, after its proposal for a "Brandt commission" on the environment took a battering at a series of environmental meetings in Nairobi.

Japan had proposed that a group of independent and widely respected personalities get together to produce a weighty document on the global environment to the year 2000. Much to the delegation's chagrin, however, the Group of 77, representing developing nations, proved solidly opposed to the idea. The group wanted direct development aid — for re-afforestation, for example — and not simply another expensive report for the bookshelves.

This was only one of the little battles which have coloured Nairobi life in recent weeks, during the three meetings which celebrated the ten-year anniversary of the Stockholm environment conference. The meetings held were one for non-governmental organizations concerned with the environment, a "meeting of special character" convened by the United Nations Environment Programme (UNEP) to review the decade, and the general council of UNEP itself.

Environmentalists have left the meetings with mixed feelings. A lot of hard political and scientific lessons had been learned over the decade, said one, but there was still a consummate lack of will to do anything practical.

The current governmental attitude was represented by the level of the delegations sent by most countries to the "meeting of special character". UNEP invited all heads of state, but in the end there were just three — Presidents Mobutu (of Zaire), Nemeiry (Sudan) and Arap Moi (Kenya). The United States sent its head of the Environmental Protection Agency (EPA), Anne M. Gorsuch, who made a fine speech recording US environmental support in the 1970s. Ms Gorsuch has been generally regarded as President Reagan's axewoman at EPA, so her speech came as a surprise. However, she left her cutting comments for a press conference — the US contribution to UNEP would fall from \$7 million in 1982 to \$3 million next year, she said. James Buckley of the State Department added that the US government saw UNEP as a "catalyst" rather than a prime actor.

One bonus for UNEP, though, came from a surprising quarter. Britain's junior environment minister Tom King promised an increase in the British contribution from £600,000 this year to £750,000 next. Even allowing for inflation, this is a substantial increase, which the UK Department of the Environment must find within its cash-limited budget. "We are conscious that resources are strained", said Mr King, and "now is the time for action".

Libya also promised its first ever contribution to UNEP — \$1 million. The Netherlands offered another 50 per cent; Japan, Finland, Malaysia, Uganda and Thailand also promised more. But these increases would not cover the big cuts threatened by the United States.

There was agreement on one thing, however: a declaration, now to be known as the "Nairobi declaration", sixteen pages of fine prose. "The world community of States solemnly reaffirms its commitment" says the declaration "to the Stockholm Declaration and Action Plan . . . It also reaffirms its support for strengthening UNEP as the major catalytic instrument for global environment cooperation . . .". The word "catalytic" is to be noted.

Robert Walgate

Information technology

Europe wakes up

Brussels

Europe's leading electronics and information technology companies are taking seriously the European Commission's grandiose plans to pool all their research efforts. Dubbed "Esprit", the European Strategic Research Programme in Information Technology, it has evolved in a series of meetings that the Commission has held with leading European electronics firms during the past year.

The first fruits of these discussions have been leaked from a communication sent to the Council of Ministers in preparation for the EEC's Science Council to be held on 30 June. Yet despite industry approval, it is in the political arena that the real fight will take place to get the Commission's ambitious ideas implemented. Discussions with UK government bodies such as the Department of Industry and the Science and Engineering Research Council have left officials in Brussels with the impression that although there is token acceptance of Esprit's principles, the national officials doubt that they can ever be realized.

Europe is losing out to Japan and the United States in the race to develop information technology, argues the Commission, not for want of spending vast sums on research. Siemens alone devotes around \$800 million a year to research. Yet Siemens, ICL and the others are in financial difficulties and failing to reap the rewards from their research investments.

Industrialists feel strongly that to get the best out of Europe's research expenditure a new body should be set up to coordinate activities. Far from wishing to build empires, Commission officials dismiss the idea that they themselves should tackle this. A more professional body such as the United Kingdom's National Research and Development Corporation with experience in turning research into marketable products would be given the job. It would be co-financed by the companies involved

and the member states, and topped up from the EEC's budget.

The management of Esprit would be decentralized and controlled by industry. The aim would be to copy the Japanese by setting clear strategic objectives and concentrating research and exploitation on a few large companies. Areas already singled out for concentrated exploitation are: advanced microelectronics; advanced information processing systems usable by the man in the street; systems for office automation and computer integrated flexible manufacturing; and fourth and fifth generation computers. The Commission is also asking for some \$11 million to spend on preliminary pilot projects to sort out the best ways of implementing cooperation.

Radical though the scheme is, the Commission is pressing for an agreement by December 1982 and a start in January 1984. Without such a European effort being launched quickly, officials in Brussels are saying, European companies will increasingly be forced to take minor roles in cooperative deals with Japanese or American competitors — in the United Kingdom ICL is already moving in this direction.

While the Commission may be accused of being over-ambitious, it is clear that Esprit is being taken very seriously by Plessey, GEC, ICL, Nixdorf, Siemens, AEG, Honeywell-Bull, CIT Alcatel, Thomson-CSF and Philips. **Jasper Becker**

Space research

UK goes national

For the first time in ten years, Britain is trying to devise a coherent space policy. The impetus has come from the recognition, later than in other Western countries, that there is money to be made from selling space technologies and services. The British government's apparently sudden decisions, earlier this year, to allow and even encourage direct broadcasting by satellite and cable television point to its fear of losing out on new technologies now coming of age elsewhere in Europe and which are already well established in the United States. The plans now being discussed for a national remote sensing programme, based on the European Space Agency's recently approved Earth Resources Satellite 1 (ERS1) are thus meant to prepare British industry for the time when money can be made from remote sensing data and hardware.

But if the spirit has changed, where will the money come from? Public money will not flow freely, so what about private investment? A select group of financiers and representatives from the space community met last week in the Surrey country house belonging to Logica, the software company, to address that question. Disappointingly the sound of

money changing hands was muffled.

What little space policy Britain has had has been directed through its membership of the European Space Agency at the space segment, the hardware in orbit. Now it seems to be agreed that more attention must be paid to the use of data sent back from space — and to persuading people that they must pay for it.

Financiers, still wary of the risks involved in satellites, see more scope for helping to market space on the ground.



Based on the experiences of direct broadcasting and cable in the United States, they are convinced that there is a market in Britain, while industry seems keen to produce antennae and dishes for receiving direct satellite broadcasts. What the financiers want to know is how big the market will be.

Remote sensing, for which the market is diffuse and ill-defined, faces different problems. A programme of scientific research using ERS1 data has been worked out by the Natural Environment Research Council and the Science and Engineering Research Council, but nobody yet knows who will pay for remote sensing data. According to Sir Hermann Bondi, chairman of the Natural Environment Research Council, Britain is fortunate that ERS1 will be an oceanographic monitoring satellite. The need to analyse ocean data quickly because its usefulness is often only immediate will, he says, pave the way for the later use of land data.

Land remote sensing data may be more saleable, perhaps to mining companies and agriculture, but only in the long run. The experience of the National Aeronautics and Space Administration in running Landsat is, however, unhelpful; the early promises made for the data have not been fulfilled. The meeting last week seemed, however, confident that data from Landsat D, to be launched later this year, and from Spot, the French national remote sensing satellite, will be more attractive.

One problem that perplexed the meeting is that the applied science of remote sensing falls administratively between two stools, the Science and Engineering Research Council (science) and the Department of Industry (application). This is why there has been a spate of interest in the question whether Britain should have a space agency of its own. Given departmental jealousy, the outcome will probably be some kind of coordinating agency, more modest, for example, than the French.

Judy Redfearn

British radioastronomy

Fear of flying

The British propensity to jetset may cause acute problems for Cambridge radio-astronomers. In the next few months Graham Eyre, QC, the inspector at the public enquiry into the proposal to site the third London airport at Stansted in Essex, will hear evidence from the University of Cambridge that an expanded airport at Stansted would seriously impede work at Mullard Radio Astronomy Observatory.

The Mullard Observatory, which is situated some 35 miles from Stansted, has for many years had an agreement with the Home Office whereby aircraft interference is controlled. Interference occurs when ground transmissions are reflected back into the radio telescopes and when aircraft



In-line for interference — the 5-km radio telescope.

equipment gives out radio waves on the telescope's frequency.

If Stansted were to be expanded to take 15 million passengers a year, the option favoured by the British government, the increase in air traffic would mean that aircraft approaching the airport would start their landing manoeuvres within 4 to 5 km of the observatory. Work using low frequency wavebands — observations of the solar and interplanetary medium and of supernova remnants, for example — could be completely disrupted.

In support of its evidence to the enquiry, the observatory will also present a technical document assessing the amount of interference and the consequences for research, which are not easily quantified. The problem is that the frequency bands allocated to radioastronomers are also shared by other radio users, at present not allowed within 50 km of the observatory. With aircraft, the range would need to be expanded to 200 km, which is not a practical option for a busy airport.

The observatory has occupied its present site since 1956 and the area, which was selected for its lack of radio interference, has been kept "clean" ever since. It would be expensive and difficult to find another suitable site.

Jane Wynn

Kibbutz biotechnology

Farming today

Rehovot

The traditional view of the Israeli kibbutz — hard work in the fields under a blazing sun — may have to be changed. For the future of many kibbutzim may lie with small science-based industry using the techniques that are already spawning new biotechnology industries in the United States and Europe.

The kibbutz at Beit Haemek is typical of the trend from farming village to light manufacturing to science-based industry. In the mid-1970s, Beit Haemek was looking for a way of supplementing its income from the cultivation of avocados and the raising of cows. And when it failed in an attempt to balance its budget by building and selling pipe organs, the kibbutz decided to establish a company that would use advanced tissue culture techniques to propagate ornamental, specific pathogen-free plants. This move made sense as the kibbutz had several members with degrees in agriculture and microbiology, and others with scientific training and laboratory experience.

Further expansion of the enterprise, now called Biological Industries, was facilitated by the arrival of a new American member in what had been a predominantly British settlement. He is Dr Robert Levin, previously a research associate at Stanford University Medical School.

The kibbutz can do only a limited amount of research on its own, and so Levin works closely with Professor Jonathan Gressel and his group at the Weizmann Institute of Science. One result of this collaboration has been the exploitation at Beit Haemek of a new Weizmann Institute system to use plant tissue culture methods for rapid and inexpensive prescreening of potential herbicides. And in cooperation with scientists at other Israeli research centres, Levin and his colleagues are now utilizing such methods to develop plant varieties resistant to herbicides, diseases and other environmental hazards.

Each successful new project has brought an expansion of Biological Industries and so now Beit Haemek has caused some surprise in Israel by placing advertisements in the newspapers inviting biologists with PhDs to join the kibbutz and applicants are now being screened. The introduction of such specialists would be quite a change in traditional kibbutz practice. On the one hand, it will be necessary for veteran members of the kibbutz to accept the fact that the newcomers will not be available, as kibbutz tradition decrees, to do any job necessary; they will expect to work in their professions. On the other hand, the new members themselves will have to come to terms with a society where there are no material rewards for their degrees or for their specific contributions.

Yet the egalitarian economic set-up of a kibbutz does seem to be consistent with making a success of a sophisticated science-based industry. Beit Haemek's enterprise expects a \$1 million turnover this year, with 65 per cent of its tissue-culture-derived products being exported, mainly to the Netherlands and West Germany.

The kibbutzniks see themselves as a production arm of Israeli academics, in the words of Beit Haemek's general manager Mike Landis, "as a conduit between Israeli research centres and the world market".

KIBBUTZ BEIT HA'EMEK
in the Western Galilee
is interest in absorbing

SCIENTISTS

at the Phd. level in the following fields:
Physiology (animal and plant),
Biology,
Agriculture,
Biochemistry and Microbiology.
A fine way to combine quality of life with
interesting work.

Please contact: The Absorption Dept., Kibbutz Beit
Ha'emek, D.N. Ma'alot Hagalit 25115.

Robert Levin of Beit Haemek sees a great deal of similarity between the science-based industries on kibbutzim and the new biotechnology companies in the United States, "which are", he says, "really cooperatives created by a few talented scientists. The only difference is that kibbutz scientists in Israel share modest living quarters while capitalistic scientists in the States share enormous profits".

Nechemia Meyers

Nature in Poland

Many readers will have been moved to ask how they can help by last week's report of the difficulties encountered by Polish scientists in acquiring scientific journals from abroad (see *Nature* 27 May, p.259). The chronic shortage of hard currency, which has apparently worsened since the introduction of martial law last December, has made the purchase of journals virtually impossible.

In the belief that many scientists outside Poland may wish to help colleagues thus deprived, *Nature* is introducing a scheme whereby scientists outside Poland may purchase a year's subscription to *Nature* in Poland at roughly half the usual price.

Those wishing to buy a subscription for a Polish colleague should write to *Nature* in either London or New York (the addresses are given opposite page 349), enclosing a cheque made payable to *Nature* for £50, or the equivalent in any currency, together with their full name and address and those of the intended recipient. Telephone orders may be charged to international credit cards. *Nature* will meet the extra cost of despatching the journal to Poland.

Immediately on receipt of an order, the intended recipient will receive a letter from the Editor describing the arrangement that has been made and identifying the donor. As soon as the order has been processed, copies of *Nature* should arrive in Poland within a week of publication. This offer will remain open at least until 1 August.

Grenoble neutron beam reactor

Working again

Europe's high flux neutron beam reactor, the Institut Laue-Langevin (ILL) at Grenoble, is likely to re-start in June with an unrepaired hole in its primary cooling circuit.

Admittedly, the hole is inside the heavy water primary cooling circuit, and is not considered dangerous, but careful experiments and calculations are being made to make sure that the reactor will work safely despite the fault.

The broken part is a section of perforated aluminium alloy, designed to break up the flow of the cooling water before it turns a corner in the piping. A 250 cm² section broke away, and parts of this interfered with the operation of a cooling pump. As soon as the fault was discovered, the reactor was shut down, and already nearly a month's experiments have been lost — a lot of work at ILL, which services dozens of visiting groups working simultaneously. The groups use the neutrons, often for only a few days, but the data so taken provide much more work on data analysis back at the group's home institutions.

Dr Brian Fender, ILL's British deputy director, said last week that he expected to be able to save the 1982 experimental programme by putting off a scheduled development period, due in November, until next February. Ultimately, some two months' experimental time at ILL will have been lost; but it will be somewhat less painful if the loss is taken next year, since 1983 experiments are still only in the planning stage.

The aluminium part broke because of ten years' cycling in turbulent water flow, ILL engineers believe. Well away from the reactor's fuel rod, it did not suffer significant radiation damage. Almost all the pieces have now been recovered. Reassembly of the jigsaw has shown a piece a centimetre across still to be missing, presumed lodged in the cooling circuits, but it is not expected to do more damage when the reactor is restarted.

Flow through the cooling circuit has been tested, using a dummy fuel element, and despite the hole in the aluminium section, no increased turbulence or vibration has been observed. The broken part has also been inspected by TV camera, from a distance of one or two metres, and seems firm enough. It is conceivable that the French licensing authorities would not allow ILL to restart without replacing the part, but, says Fender, "we work very closely together", and he does not expect any trouble from that quarter.

Why could the broken part not simply be replaced? Because it is awkward to reach, requiring the dismantling of most of the reactor, an exercise which could cost many more months than the present delay.

Robert Walgate

CORRESPONDENCE

A letter to Soviet scientists

The following document reached *Nature* on 21 May 1981, the 61st birthday of Academician Andrei Sakharov.

I have learned from foreign broadcasts that the "Scientists for Sakharov, Orlov and Shcharanskii" (SOS) Committee in the United States has issued an appeal to Soviet scientists to come to the defence of the family of Sergei Kovalev — that is, Sergei himself, his daughter-in-law Tat'yana Osipova, and his son Ivan. I feel sure that many of you also heard this appeal. I beg you to treat it with all seriousness and to undertake those actions to which each of you feels prompted by his feeling of responsibility as a citizen and by his conscience, both in this particular case, and in others where your intervention would be of special importance.

For many years now, I have observed an almost total lack of action on the part of my Soviet scientific colleagues in matters relating to the defence of human rights. It is shameful that foreign scientists (and deepest thanks to them!) show a greater concern for our affairs than we do ourselves. In this letter, I am speaking first and foremost to those scientists occupying a fairly high and independent position — more specifically to Members and Corresponding Members of the Academies. I know that some of you have negative or sceptical feelings regarding my statements on general questions. But this is not a general question (on these, I have always expressed my views in the form of a discussion, and have never thrust my opinions on to anyone). This is a matter of the fate of real people, including the fate of your colleague in science. Every one of you who has shown even the least interest will form his or her own independent opinions on these cases, of the total degree of injustice and the harshness of repression.

You cannot consider that these cases do not touch you; the historical experience of our country, professional solidarity, simple human sympathy for the fate of others, and often personal contact with the victims of repression (many of you were personally acquainted with Orlov, Kovalev and others), make such an attitude impossible. You must not invoke the interests of your job, the need to keep your professional post in the interests of science. In fact, most of you are in a fairly secure position. And the interests of science must include the defence of members of the scientific community from injustice; they include the responsibility of the citizen. This is not the Stalin era; practically speaking none of you *now* are under any kind of threat. But can you rule out completely a return to a new era of mass repression? For if this should come to pass, your lack of action *now* will be one of the causes. And conversely, could not the activity as citizens and the independence of even a few of the country's top scientists have a profound and beneficial effect on the whole situation?

And which of you can lay his hand on his heart and proclaim that this is not a vital necessity? It is precisely this which more than anything would promote international scientific relations, and, on a far far wider scale, trust and peace throughout the world.

Such is the measure of the individual personal responsibility of every one of you.

In conclusion, I should like to cite several examples of scientists and scholars who are victims of repression — prisoners of conscience. For each of these cases there is information which is accessible to you.

- (1) The case of Kovalev and his family, with which I began this letter. Kovalev has served a term of seven years imprisonment under the harshest conditions, and he is now in exile, under shocking conditions in the Magadan region. Tanya Osipova and Ivan Kovalev are in prison, and it is necessary to obtain for them, at the very least, the opportunity of meeting. After the arrest of his wife, and before his own arrest, Ivan had only one 30-minute meeting with her in 15 months.
- (2) The case of Yuri Orlov. His fate, and the cruel torments he is suffering are widely known.
- (3) The case of Anatolii Shcharanskii — also very well known.
- (4) The case of the mathematician and cyberneticist, Dr Aleksandr Bolonkin. After imprisonment and exile which robbed him of nine years of his life, beatings-up, confinement in the punishment cells and provocations, he is back in prison again on a trumped-up criminal charge.
- (5) The case of the philologist and poet, Vasyl Stus'. After many years in prison and exile, he was condemned to a further 15 years labour-camp and exile, simply for supporting the Ukrainian Helsinki Watch Group.
- (6) The case of Anatoliy Marchenko. He is not a scholar; he is a worker and writer, the author of some remarkable books on present-day labour camps and prisons. But he has been sentenced again — his fifth sentence — to the monstrous term of 15 years imprisonment and five years exile; the central point of the charge being his open and noble-hearted letter on the "Sakharov case" to academician P. L. Kapitsa, who is one of those to whom I address this letter.
- (7) The cases of Aleksandr Lavut and Tat'yana Velikanova (mathematicians), Mart Niklus (ornithologist), Leonard Ternovskii (radiologist), Viktor Nekipelov (a pharmacist and talented poet), Pyatkus (philologist), Meilanov (mathematician), Luk'yanenko (jurist), Airikyan, Altunyan and many others. None of them has had recourse to violence nor would do so. All of them have suffered cruelly for loyalty to their noble convictions, the principles of free expression and justice. Help them — this is our common duty.

Andrei Sakharov, *Gor'kii*, 30 March 1982.

POSTSCRIPT — In defence of Ivan Kovalev.

Ivan Kovalev has been sentenced. This is yet another blow to the champions of human rights in the Soviet Union, to free expression, and to the Helsinki Watch Group; a blow against the principles expressed in the Helsinki Final Act. And this is a new act of inconceivable cruelty towards the famous Kovalev family, the third of its victims. In December 1974, Sergei Kovalev, Ivan's father, was arrested. In May, 1980, Tat'yana Osipova, Ivan's wife was arrested. Both of them, like Ivan Kovalev himself, have received

cruel sentences for their nonviolent defence of human rights, for their fidelity to noble convictions. In a letter which he wrote shortly before his arrest, Ivan Kovalev wrote about himself and his situation. And in his article "My Tanya", he told us about his wife, her fate and her struggle. There is little one can add to those documents — and having known Ivan for many years I am convinced of the justice of his case.

I call upon everyone who believes in justice, who can feel for the tragedy of this young family, to speak out in defence of Ivan Kovalev, his wife and, of course, his father. The possibility of a meeting between Tat'yana and Ivan must be obtained — this is their inalienable human right. And it is very important too to obtain their freedom, and the freedom of all prisoners of conscience.

Andrei Sakharov, *Gor'kii*, 2 April 1982.

ESA clarification

SIR — May I refer to your article "Antarctic research hit by crisis" published in *Nature* of 15 April (p.593) and in particular to the part of the article:

"Although Argentina's own research efforts have been extremely limited, it is at present collaborating with the French in glaciological work, and plans exist to site a ground station for the first European Space Agency remote sensing satellite to be put into polar orbit, at the Argentinian base of Marambio".

There is no Argentinian connection with the European Space Agency/ERS-1 programme. There will certainly be some ERS-1 overseas ground-stations outside Europe, but no negotiations have been made with Argentina or indeed with any other South American state. If such plans were ever envisaged, they would require approval of ESA Council.

You will appreciate ESA's desire to have this point clarified.

W. BRADO
(Head of ESA Director
General's Cabinet)

ESA, Paris, France

Editor in false garb

SIR — If both of your bizarre and intemperate recent leading articles ("Doves in false garb", 18 February, p.542; and "Professional propaganda", 1 April, p.380) had been published on 1 April, one might more easily have understood them. They are so full of rancid sarcasm, repetition to the point of perseveration, wild claims devoid of any supportive evidence, multiple errors in basic logic, frank contradictions and odd grammar, that they imply a writer almost incoherent with anger. They are also probably the most disgraceful abuses of editorial privilege I have ever seen in print.

You have used, however ineffectually, the anonymity of the editorial and the prestige of a hitherto distinguished scientific journal, to lend spurious respectability and authority to your personal and highly idiosyncratic views. You have jeered at a very large number of sincere, ethical, distinguished and honourable professionals throughout the world, impugning their honesty, without raising a

single justified or logical argument. You misrepresent the views and claims of those you disagree with, so as to more easily oppose them.

You criticize the professions (with biased inaccuracy) for ignoring public issues, and belittle them when they speak out on them. You graciously grant us the responsibility to contribute our special knowledge to the solution of such problems, but criticize us for failing to exceed the limits of such special knowledge. You insist that it's laudable for us to express ourselves singly, but dreadful to do so collectively. You sneer at the "catchy acronymic titles" of the groups concerned, though few of them are catchy, and none are acronyms. You complain that the groups "claim professional support" in a deceitful way. While the degree of professional support they have is significant and growing, their titles very specifically and accurately state their membership as limited to those members of each respective profession who are opposed to nuclear arms or war: how much more accurately could they describe themselves?

Take your own advice, Mr Editor. Put aside your pique about the professions you don't belong to; put aside your anonymity and the authority that belongs to *Nature* and not to you — take your courage in your hands as you exhort us to do (and as we have done) and join us on the hustings. If you can express your views coherently and support them with data, we will listen with interest and respect.

Amongst your hectic hubris, you graciously concede that "it would . . . be shocking if physicians were silent about the probable consequences of smoking". May we not also warn about the dangers of being smoked?

MICHAEL A. SIMPSON

Health Sciences Center,
Temple University,
Philadelphia, Pennsylvania, USA

Ab what?

SIR — In the preamble to the first editorial in the edition of *Nature* dated 22 April (p.693), the word "abreact" is used. So far as I can discover the word is a term from psychoanalysis referring to the liberation of the individual from neurosis by the expression of repressed emotion. Is this what the leader article intends?

I feel it would be a pity if *Nature* were to start falling into the trap of employing jargon terms made meaningless by fashionable use.

It might cause some adverse reaction among the readers.

ROBIN H. C. STRANG

Department of Biochemistry,
University of Glasgow,
Glasgow, UK

Data protection

SIR — I read with great interest the article "Big Brother's law" (*Nature* 1 April, p.694) about personal data protection. I would like to add a few comments.

In the summer of 1981, the European Commission addressed a formal recommendation to European Community member states to sign and ratify the convention of the Council of Europe by the end of 1982. Up to now, only five member states have signed it.

In this recommendation, the Commission

stated that it would draft a Community directive if the member states failed to sign the convention.

The socialist group in the European parliament has long pressed for a European Community directive. Indeed, in March 1982 the European parliament adopted a report of its legal committee (the report was given by Herr Sieglerschmidt, SPD) asking among other things for a Community directive for personal data control.

We believe, however, that the most important point is the control of the transnational data transfer of private companies, although you are right in the statement that the transfer of information between governmental departments is very important.

Finally, I would like to say how important it is to lobby the British members of the European parliament. The international companies are very familiar with effective lobbying of MEPs. The electorate should do the same, but in the interest of maximum protection of data on individuals.

FRITZ GAUTIER, MEP

Braunschweig, FRG

Patent sense

SIR — The report in *Nature* of 29 April (p.792) of the 21 April Colloquium at the Association of the Bar of the City of New York is inaccurate in stating that a European visitor may, under European laws, derive an invention from a seminar or talk in America and appropriately apply for a patent in Europe.

European and other patent laws do not permit patents to those who derive. Typically, they require that the subject of a patent be the original work of the applicant or its predecessor in interest.

The issue of whether or not an invention is original work is different from the issue of whether or not the inventor's (original) work preceded the work of another.

DAVID W. PLANT

Fish & Neave,
New York, USA

Ageing controversy

SIR — I am writing to vehemently protest about the recent *News and Views* article from your ageing correspondent in the 1 April issue of *Nature* (296, 392-393; 1982). The work of Dr Obispo and his colleagues on "longevin", the lifespan-extending protein from carp gut is described in excruciating detail, but no mention at all is made of Dr A. Huxley's pioneering work.

It was in the obscurity of his Oxford college rooms that the re-discovery of the rejuvenating effect of raw fish viscera was made (*Phil. Trans. R. Soc. B* 240, 153-215, 1969). Furthermore, he has gone a long way in the characterization of the same protein. I am truly surprised that this work of a British scientist (who in addition comes from a very distinguished family indeed) should be overlooked in the pages of *Nature*. I think the reason for this is not only that the Obispo and Maunciple group has essentially unlimited funding from the Stoyte Foundation for Aging Research, but also that nowadays anything having to do with gene cloning receives

immediate attention, while people like your correspondent tend to overlook standard protein chemistry work. It is sad that Dr Huxley's work has been ignored for so many years, but it is even more sad that the California group should be allowed to patent the longevin B-chain clone. The extension of human lifespan is of utmost ethical importance and should not be treated lightly; least of all in a patent office.

AN AGEING PROFESSOR

Basle, Switzerland

SIR — I read the *News and Views* article on longevin (*Nature* 296, 392-393; 1982) with keen interest but think, if I may say so, that this may prove to be the swan song of geriatric biochemistry.

JENNIFER REED

Institute for Experimental Pathology,
German Cancer Research Center,
Heidelberg, FRG

SIR — Huxley in his seminal work *After Many a Summer* (Chatto and Windus, 1937) was the first to draw wider attention to the research of Obispo. It is regrettable that your correspondent's otherwise excellent report on anti-senescence factors (*Nature* 296, 392-393; 1982) did not refer to Obispo's pioneering studies, even though these were conducted on whole carp *flambé* (*J. Carp Sci.* 1, 69; 1936). Hauberk's results on raw carp viscera were not rediscovered until the regular visits of a member of the present shadow cabinet to Tashkent led to the realization that the famous mad ape in the local zoo was none other than the now fully evolved Lenin. Obispo himself, his procreative and creative powers equally unimpaired, will reach the age of 100 in 1984, by which time, we can be confident, gerontology will have experienced a timely death.

RICHARD GLENDALE

London W6, UK

Nuclear options

SIR — David Fishlock should keep his journalistic innuendos (*Nature* 22 April, p.700) for *Financial Times*. He implies that my letter in *Nature* of 18 March (p.192) on "Electricity costs" is equivalent to flat-Earth theory, ostensibly because it cites a paper of mine¹ as "in the press". He calls this "unpublished", but in fact Energy Policy No. 2, containing the paper, was published on 5 May, and the *Financial Times* was sent a copy of the page proofs before publication. In any case the scientific editor of the *Financial Times* was surely not unaware of the report of Sir Kelvin Spencer's committee², which was given wide publicity after a press conference in the House of Commons on 2 February, and which quoted or paraphrased the main results of the "unpublished" paper.

I can only assume that any conclusion which does not wholeheartedly support the nuclear case is necessarily flat-Earth theory to David Fishlock, and need not be looked at any further.

J. W. JEFFERY

Department of Crystallography,
Birkbeck College,
London WC1, UK

1. *The Real Cost of Nuclear Electricity in the U.K.*, Energy Policy, Vol. 10, No. 2, 76-100 (1982).
2. *Nuclear Energy: The Real Costs: A Special Report by the Committee for the Study of the Economics of Nuclear Electricity*. (Worthyvale Manor, Camelford, Cornwall).

NEWS AND VIEWS

Gravitational radiation and the binary pulsar

from Virginia Trimble

GRAVITATIONAL radiation not only exists, but it is apparently also doing exactly what general relativity says it ought to, at least in the binary pulsar PSR1913 + 16. This is the conclusion drawn by J.H. Taylor and J.M. Weisberg writing recently in the *Astrophysical Journal* (253, 908; 1982). It comes after more than six years of observing gradual changes in the eight-hour orbit of the pulsar and its invisible companion, probably also a neutron star.

The pulsar period itself (0.059 s) acts as the known emitted frequency from whose Doppler shifts the orbit is determined as part of a total solution including up to 20 variables. These represent intrinsic properties of the pulsar, classical orbital elements, and relativistic corrections for the advance of the perihelion (about 4° per yr compared with $43''$ per century for Mercury) and variable parts of the gravitational redshift and the transverse Doppler effect. The relativistic parameters together define a small set of allowed values for the masses of the two stars. These, plus the orbit period, can be plugged into a standard formula for gravitational radiation from orbiting point masses to yield a unique prediction of the rate at which the orbit should be decaying, $\dot{P} = -2.4 \times 10^{-12}$. The measured value, from the multi-variable fit, is $-2.30 \pm 0.22 \times 10^{-12}$, in exact agreement to within the experimental errors.

The high precision and self-consistency of these results mean that they have interesting implications for three problems of current interest in astrophysics.

First, if no important effects have been left out, the binary pulsar orbit provides as accurate a mass for a neutron star as has ever been determined, and the first one for an honest-to-goodness radi pulsar. The other measured masses all come from X-ray binaries, for which the energy source is not the rotation of the neutron star (as in 'true' pulsars) but accretion of gas from a companion, an inherently messy process. The answer is $1.4 \pm 0.1 M_\odot$ for each component. Existing data¹ are² consistent with all known neutron stars having about this mass. It is just above the maximum stable mass (Chandrasekhar limit) for a star or stellar

core supported by degenerate electron pressure and may reflect a fairly universal truth — you get a neutron star when an inert core builds up to the Chandrasekhar limit and so collapses.

Second, we are encouraged to ask more seriously than before how gravitational radiation will affect the evolution of other short-period binary systems, whose orbits are less clean than that of PSR1913 + 16, but which are astronomically at least as interesting. The classic case is that of the cataclysmic variables (CVs)²⁻⁶ — novae and related systems consisting of a normal star transferring material to a white dwarf companion. Many kinds of binaries show such mass transfer. In some, the flow is driven from a more massive to a less massive star by newtonian effects; in others, a star spills over because it is trying to expand and become a red giant⁷. But in some low-mass CVs, neither of these explanations works. The shedding star is the less massive and has not yet begun to expand. What drives the transfer? Gravitational radiation must gradually bring the two stars closer together. The surprise is that it does so at precisely the rate needed to produce the otherwise inexplicable mass transfer in these short-period, low-mass CVs⁸⁻¹⁰. In addition, beyond a certain point, gravitational radiation-driven mass transfer begins to move the stars apart again, and so may also explain some puzzling gaps in the distribution of orbit periods of CVs.

Third, the binary pulsar provides a test of the standard formula, first derived by Einstein¹¹, for the gravitational radiation expected from a binary system. Over the years, many theorists have tackled the problem within the framework of general relativity. Some have confirmed the so-called quadrupole formula¹²⁻¹⁴, whilst others have found different results, or, more often, concluded that a proper derivation has not yet been carried out and is currently beyond them, too¹⁴⁻¹⁶. The agreement to within 10 per cent between

prediction and observation for the binary pulsar strongly suggests that the quadrupole formula is very nearly right, at least for relatively weak-field cases (again, unless some important effect has been left out of the multi-variable analysis). This should help theoretical relativity both with this particular problem (it is always easier to do a calculation if you know the answer in advance) and with others by suggesting appropriate approximation methods. General relativity is not the only theory of gravitation currently under investigation. But most of the others (except some cases of the Brans-Dicke theory) predict orbit changes due to gravitational radiation that disagree with the binary pulsar result¹⁷. We cannot, however, quite rule them out yet, if only because the calculations leading to that conclusion have been done with methods analogous to those that lead to the quadrupole formula for general relativity, which may, just possibly, not be appropriate.

What can we expect from PSR1913 + 16 and gravitational radiation in the future? Observations are continuing. And the nature of the problem is such that precision increases roughly with the square of the observing time. Thus another five years could narrow the error bars on the relativity parameters from about 10 per cent to about 1 per cent, providing neutron star masses with that precision and testing agreement (or disagreement) to 1 per cent between the quadrupole formula prediction and the observed value for the orbit decay due to gravitational radiation. Beyond this we probably cannot go. Fluctuations in the newtonian gravitational potential of the Galaxy (due to stars, clusters, gas clouds, spiral arms and all the rest) cause random, undecipherable accelerations of a test particle and so will put noise into the pulsar timing measurements at a level near 2×10^{-14} on time scales comparable with the effects being sought.

Finally, once a phenomenon has shown up indirectly in an astronomical context, it is natural to ask whether it can be seen directly in the laboratory. Detectors designed to search for gravitational radiation pulses have been in fairly continuous operation one place or another for

Virginia Trimble is in the Department of Physics, University of California, Irvine, California 92717, and the Astronomy Program, University of Maryland, College Park, Maryland 20742.

about 15 years. Ferrari *et al.*¹⁸ have reported a recent data set in which widely separated detectors were simultaneously excited by some sort of background and reference earlier results. Helium-cooled detectors of higher sensitivity are being developed or are in operation at Rome, Stanford, Maryland, Tokyo and other places. The expected gravitational

radiation flux at the Earth from the binary pulsar is, however, only about 10^{-8} ergs $\text{cm}^{-2} \text{s}^{-1}$, less than 10 per cent that expected from the Crab Nebula pulsar, and some 10 orders below the sensitivity of foreseeable ground-based equipment at the relevant frequencies. Indirect methods will, therefore, clearly continue to be important for a long while, at least for PSR1913 + 16!

1. Bahcall, J.N. *Astr. Astrophys.* **16**, 241 (1978).
2. Kraft, R.P., Mathews, J. & Greenstein, J.L. *Astrophys. J.* **136**, 312 (1962).
3. Braginskii, V.B. *Usp. Fiz. Nauk* **86**, 433 (1965).
4. Paczynski, B. *Acta astr.* **17**, 207 (1967).
5. Faulkner, J. *Astrophys. J. Lett.* **170**, L99.
6. Pringle, J. & Webbink, R.F. *Mon. Not. R. Astr. Soc.* **172**, 493 (1975).
7. Paczynski, B. *Astr. Astrophys.* **9**, 183 (1971).
8. Paczynski, B. & Sienkiewicz, R. *Astrophys. J.* **248**, L27 (1981).
9. Paczynski, B. & Sienkiewicz, R. *Preprint* (Orange Aid Series, 1982).
10. Rappaport, S., Joss, P.C. & Webbink, R.F. *Astrophys. J.* **254**, 616 (1982).
11. Einstein, A. *Sber. preuss. Akad. Wiss. Phys.-Math K1*, 154 (1918).
12. Landau, L. & Lifschitz, E. *Classical Theory of Fields* (Addison-Wesley, Massachusetts, 1951).
13. Walker, M. & Will, C.M. *Astrophys. J. Lett.* **242**, L129 (1980) and refs therein.
14. Cohen, J.M. *Ann. N.Y. Acad. Sci.* **375**, 459 (1982) and refs therein.
15. Ehlers, J., Rosenblum, A., Goldberg, J.N. & Havas, P. *Astrophys. J. Lett.* **208**, L77 (1976).
16. Cooperstock, F.I. & Hobill, D.W. *Astrophys. J. Lett.* **235**, L55 (1980).
17. Weisberg, J.M. & Taylor, J.H. *Gen. Rel. Grav.* **13**, 1 (1981).
18. Ferrari, V., Pizzella, G., Lee, M. & Weber, J. *Phys. Rev. D25* (May 15, in the press).

Antibodies and cancer therapy

from P.C.L. Beverley

A RECENT report describes the dramatically successful use of a monoclonal antibody in the treatment of a human B-cell tumour¹. The patient has now been clinically free from disease for eleven months without further treatment. This striking success was largely due to the exquisite specificity of the antibody for the tumour cells, but is something of a special case. The monoclonal antibody was anti-idiotypic: that is, it was raised against specific determinants on the antibody expressed on the surface of the tumour cells. The method would thus seem limited because a different antibody has to be produced for each individual tumour and because only B-cell tumours, which are uncommon, are known to carry sufficiently specific determinants. Whether it will be possible to find markers that will allow a similar specificity to be gained against other tumours is not yet clear.

Several other patients have, however, already been treated with monoclonal antibodies and it is possible to assess something of the general potential of antibodies for therapy.

The questions to be faced are: are the monoclonal antibodies safe, and are they effective? So far, there is a clear answer to only the first question. Administration of mouse monoclonal antibodies produces remarkably few side effects even when free antigen is present in the serum and it must be presumed that the infusion of antibody leads to the formation of circulating immune complexes¹. At first sight the apparent safety of monoclonal antibodies might be assumed to stem from their monospecificity — they do not contain the contaminating specificities that give problems when conventional antibodies, for example anti-lymphocyte globulin, are

administered. This may, however, be an over-simplification since it has been predicted that many monoclonal antibodies will give unexpected cross-reactivities². Indeed, cross-reactivity has been demonstrated for some of the antibodies which have been used *in vivo*: for example, the J5 antibody used in treatment of acute leukaemias³ was found also to react with renal tubules. The moral to be drawn is that before an antibody is used *in vivo* it is wise to screen widely for unexpected reactions with different tissues.

If the antibodies are safe the major problem of antibody specificity remains. Antibodies to differentiation antigens can be used when the loss of normal cells with the same antigens would not matter or when tumour cells have completely replaced the normal population. Thus, antibodies against T-lymphocyte antigens have been used to treat advanced stages of mature T-cell tumours. In the first patient, a significant response of skin lesions was seen and treatment was continued for 17 weeks until the tumour progressed in lymphoid organs⁴. This patient showed no immune response to the mouse immunoglobulin but in subsequent trials and in patients given anti-T cell antibody as immunosuppression for renal allograft rejection episodes, antibody to the mouse protein was produced and limited the duration of treatment⁵.

While antibodies to differentiation antigens can be used for tumour therapy, their potential would be much greater if tumour-specific antibodies were available. Several claims for the existence of such

antibodies have been made, notably of those to a common acute lymphoblastic leukaemia antigen (CALLA)⁶. However, further examination in another laboratory showed the antigen to be present on some normal bone marrow cells⁷, suggesting it might be best to treat all claims with some scepticism, at least until candidate antibodies have been examined in several laboratories. Nevertheless, an anti-CALLA antibody (J5) was used to treat leukaemic patients and revealed another problem of antibody therapy, that the tumour cells rapidly lost the target antigen when antibody was administered. Similar antigenic modulation could be demonstrated *in vitro*.

While treatment *in vivo* has not yet produced spectacular results, with the exception of the anti-idiotypic antibody, treatment *in vitro* holds more immediate promise. Bone marrow grafting has become a recognized treatment for leukaemia but has considerable problems, including graft-versus-host disease (GVH) caused by T lymphocytes in the donor marrow. Removal of T lymphocytes from bone marrow before grafting had been attempted with conventional antisera and it was a logical step to see if better results could be obtained with more specific monoclonal antibodies. So far the published data are limited to a single study in which the marrow cells were merely incubated with monoclonal anti-T cell antibody before infusion⁸. The procedure relies for effect on the removal of antibody-coated cells from the circulation by the reticuloendothelial system — hardly likely to be very effective in transplant recipients who have been irradiated and treated with cytotoxic drugs to condition them to accept the graft. Nevertheless, the early results are not discouraging and the use of antibodies with complement or coupled to drugs or toxins should improve its effectiveness. Leukaemia can also be treated by an autologous transplant. Remission bone marrow is collected and stored until the patient relapses. Aggressive chemo- and radiotherapy can then be administered and the patient rescued with the stored marrow. This approach suffers from the problem that the re-infused marrow may contain leukaemic cells but these might be removed if sufficiently specific antibodies can be produced⁹. *In vitro* treatment has the advantage that surplus antibody, or antibody-drug or antibody-toxin conjugate, can readily be washed away before the marrow is infused.

It is notable that the research discussed so far all deals with tumours of the haematopoietic system and common tumours (all carcinomas) have not been mentioned at all. This is in part because many of the earliest and best characterized monoclonal antibodies have been directed to leukocyte antigens, and also because the therapy of solid tumours poses additional problems, particularly the penetration of antibody into tumour masses. Nevertheless

P.C.L. Beverley is in the ICRF Human Tumour Immunology Group, School of Medicine, University College London, University Street, London WC1E 6JJ.

an initial trial has been made, although without great effect, in patients with extensive disease¹⁰. Once again, bone marrow clearance *in vitro* may soon be the most useful manoeuvre, at least for some solid tumours which are chemo- or radiosensitive.

What of the future? First it is clear that antibody alone is unlikely to be optimally effective and drug- or toxin-coupled antibodies may provide additional therapeutic advantage, particularly *in vitro*. *In vivo* therapy is frequently limited by the development of anti-mouse immunoglobulin antibody but it may be possible to induce tolerance or, in the future, to reduce the problem by the use of human

monoclonal antibodies. Already the development of human anti-tumour antibodies has been reported¹¹. While serotherapy will surely not be a panacea, there are grounds for cautious optimism.

1. Miller, R.A., Maloney, D.G., Warnke, R. & Levy, R. *New Engl. med. J.* **306**, 517 (1982).
2. Lane, D. & Koprowski, H. *Nature* **296**, 200 (1982).
3. Ritz, J. *et al. Blood* **58**, 141 (1981).
4. Miller, R.A., Maloney, D.G., McKillop, J. & Levy, R. *Blood* **58**, 78 (1981).
5. Cosimi, A.B. *et al. New Engl. med. J.* **305**, 308 (1981).
6. Ritz, J. *et al. Nature* **283**, 583 (1980).
7. Greaves, M.F. *et al. in Modern Trends in Human Leukaemia* Vol. 4 (eds Neth, R. *et al.*) 296 (Springer, Berlin, 1981).
8. Prentice, H.G. *et al. Lancet* **i**, 700 (1982).
9. Storb, R. *Nature* **295**, 555 (1982).
10. Sears, H.F. *et al. Lancet* **i**, 762 (1982).
11. Sikora, K. *et al. Lancet* **i**, 11 (1982).

Plant cells on the move

from Clive Lloyd

CORSETED in cellulose, plant cells have limited freedom of movement. As a result, their morphogenetic responses, rather than consisting of cellular migration, involve internal rearrangements of cytoplasm. The responses include alterations in the rates of cell division and of enlargement (which determine the overall size of cells). Fine tuning of tissue morphology is, however, largely a product of the shapes adopted by cells and of the orientation of the walls which are deposited between them at cytokinesis. A common element uniting all these processes is the microtubular cytoskeleton since microtubules: (1) help to determine the polarized shape of interphase cells by somehow guiding the deposition of cellulose; (2) as the preprophase band (PPB) they foretell where the cytokinetic cell plate will be deposited; (3) help form the mitotic spindle; and (4) contribute to the cytokinetic apparatus, the phragmoplast. The microtubule cycle can therefore be regarded as at the hub of key morphogenetic processes. To come to terms with this at the level of control it will be important to obtain a molecular description to complement the existing ultrastructural one. For instance, where are the microtubule-organizing centres located? Is there one set responsible for all microtubule arrays or are there several? If several, how are their functions co-ordinated?

In comparison with animal tissue culture cells, plant cells are much less tractable and, as a consequence, the successful application of molecular probes to the study of the cytoskeleton has been slow. But over the last year several studies have emerged which present real technical advances that should lead to a more dynamic picture of the microtubule cycle.

One possibility when considering the microtubule cycle is that different tubule arrays are composed of different tubulins

produced, perhaps, in a cycle-dependent manner. The multitubulin concept was reviewed by Fulton in a recent *News and Views* article¹ which outlined that non-mammalian organisms do contain multiple genes for tubulins and that the α -subunits of their tubulin heterodimers may differ from those of vertebrate brains, which, unlike most other tissues and unicellular organisms, constitute a rich source of microtubule proteins. Until now the lack of grey matter has held back research into plant tubulin, but in this issue of *Nature* (p.426), Morejohn and Fosket report its isolation from Paul's scarlet rose suspensions and, for the first time, its reassembly into microtubules.

Plant extracts, prepared according to methods for brain tubulin and containing leupeptin as a protease inhibitor, were fractionated on DEAE-Sephadex. One fraction was enriched in proteins which migrated on SDS gels in the vicinity of brain tubulins. Indeed, microtubules could be produced from this fraction either by self-assembly in glycerol or by 'precipitation' with the plant metabolite taxol which enhances the polymerization of tubulin². It is of interest that comparative peptide maps suggest that whereas brain and plant β -tubulins are very similar, the α -subunits are different. As Clayton *et al.* have suggested³, this could be the basis of the relative resistance of the microtubules of microorganisms to drugs such as colchicine and this explanation would now seem to apply to higher plants as well. Provided that satisfactory cell synchrony can be obtained, the stage is now set to see whether there are cycle-specific classes of plant tubulins.

Clive Lloyd is at the Unilever Research Colworth Laboratory, Sharnbrook, Bedford MK44 1LQ.

Several recent histological studies have outlined the dynamic nature of plant microtubules. De Mey *et al.*⁴ reported the staining of microtubules in the endosperm of *Haemanthus* (giant African blood lily) by an immuno-gold method. Using this technique, anti-tubulin antibody-decorated microtubules can be made visible by the subsequent addition of colloidal gold-tagged antibodies and beautiful pictures have been obtained of interphase and mitotic arrays. One point to emerge from this work is that cytoplasmic microtubules seem to be grouped around the nucleus to the extent that it is concluded "that the nuclear envelope or the nucleus, or both, is instrumental in tubule nucleation and arrangement". This is intriguing since sites of microtubule nucleation in acentriolar plants are poorly understood and difficult to pinpoint, and this study gives (some of) them a definite perinuclear location. The function of these cytoplasmic microtubules in endosperm (which, at stages, is liquid; containing wall-less nuclei in a syncytium) is enigmatic and had not been previously detected. It is therefore possible that they have a different function to the plasma membrane-associated microtubules believed to help orientate cellulose microfibrils and which have been reported to emerge, in *Azolla* roots, from nucleation sites located at the cell periphery⁵. Another interesting feature of this study is that it shows cytoplasmic microtubules co-existing with phragmoplast tubules which, in turn, overlap in time with the anaphase mitotic spindle. This implies separate microtubule-organizing centres or nucleation events for different microtubule assemblies whose bouts of activity are not mutually exclusive.

The large nuclei of *Haemanthus* endosperm make it ideal material for the study of mitosis and, as Euteneur and McIntosh⁶ have shown with their work on microtubule polarity, for phragmoplast formation also. Interphase cytoplasmic hoops of microtubules are absent from this tissue, however, and other material must be used to study them. The difficulty with applying antibody-based techniques to walled cells is that the wall generally excludes antibodies, but there are now ways around this problem. By pre-fixing roots with aldehydes, Wick *et al.*⁷ were able to preserve cell shape before degrading the wall with enzymes and with EGTA (which presumably weakened intercellular calcium pectinates). Following this treatment, gentle squashing released cuboidal cells which could at this stage be permeabilized with acetone before carrying out double immunofluorescence using anti-tubulin antibodies. For the first time, the PPB has been visualized by immunofluorescence and all four assemblies are beautifully illustrated. Some interesting aspects deserve comment. For example, at the same time as the PPB is stained, so is the nuclear envelope, which appears to have

microtubules — not of the PPB — associated with it. In some cases where the nucleus has been freed from the rest of the cell, the PPB segregates with the nucleus and not with the cell cortex.

It is early days in this more detailed phase of plant cell biology but already the new approaches are presenting tantalizing and unexpected findings. With such techniques now successfully applied to higher plants,

we can expect to learn more about the spatio-temporal aspects of microtubule nucleation which accompany key processes in plant morphogenesis. □

1. Fulton *Nature* **296**, 308 (1982).
2. Schiff *et al.* *Nature* **277**, 665 (1979).
3. Clayton *et al.* *FEBS Lett.* **115**, 301 (1980).
4. De Mey *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 1898 (1982).
5. Gunning *et al.* *Planta* **143**, 161 (1978).
6. Euteneur & McIntosh *J. Cell Biol.* **87**, 509 (1980).
7. Wick *et al.* *J. Cell Biol.* **89**, 685 (1981).

Electrons in novel two-dimensional structures

from J.M. Worlock

THE report that Gottfried H. Döhler and his colleagues have been able to grow a specimen of gallium arsenide (GaAs) in the form of a doping superlattice¹ is an exciting landmark in semiconductor physics. For Döhler, primarily a theoretician, at the Max-Planck Institut für Festkörperforschung in Stuttgart, the development is the culmination of ten years' work. Having discerned the properties expected of a doping superlattice², he has now persuaded colleagues in Stuttgart and Munich to grow the new lattice and to confirm his predictions.

Interest in semiconductor superlattices goes back to 1970 when Esaki and Tsu³, at

IBM, perceived that it was possible, at least in principle, to improve on nature and design semiconductor microstructures with carefully tailored properties.

They envisioned two types of semiconductor superlattice. One, a doping superlattice, is that analysed by Döhler (and independently by Yu. A. Romanov⁴) and recently realized. The other, a compositional superlattice, was realized as early as 1974 by groups at IBM⁵ and Bell Laboratories⁶ and has been much studied⁷.

Figure 1 shows the spatial variation of the electron bands in the direction of the superlattice. The doping superlattice (Fig. 1a) is seen to those familiar with semiconductors to be simply an alternation of p-n and n-p junctions. The periodic rise and fall of the conduction and valence bands is caused by a periodic variation of impurity concentration. The rise and fall also mimics on a more macroscopic scale the periodic attraction and repulsion that electrons experience in the microscopic atomic lattice: electrons are attracted to minima in the conduction band while holes are attracted to maxima in the valence band.

In a compositional superlattice (Fig. 1b), on the other hand, alternating layers are of different elemental composition. Locally, each composition has its own band structure, so that within limits the periodic variation of the conduction and valence bands can be arbitrarily chosen. The periodic structure shown in the figure resembles the familiar case of the combination of GaAs, with bandgap E_{g1} , with an alloy semiconductor aluminium-gallium arsenide (AlGaAs) having a slightly larger bandgap. A more recent development, the type 2 superlattice, conceived at the IBM laboratories⁸, is exemplified by the combination of indium arsenide (InAs) and gallium antimonide (GaSb). Here the valence and conduction bands interpenetrate, resulting in a much more complicated superlattice.

Döhler *et al.*'s doping superlattice has

J.M. Worlock is in the Bell Telephone Laboratories, Holmdel, New Jersey 07733. A. Pinczuk collaborated in the preparation of the article, and D. Olego provided helpful criticism.

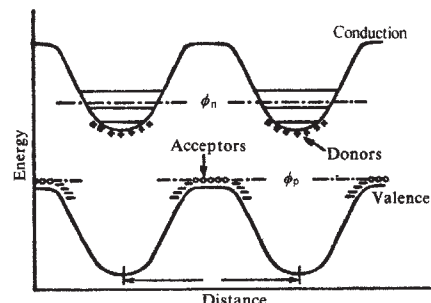
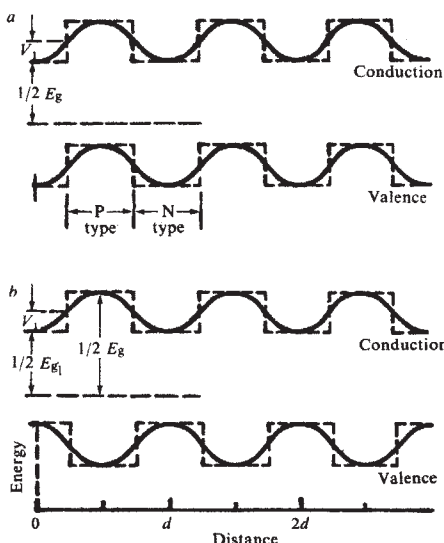


Fig. 2 Spatial variation of the conduction and valence bands in an excited state of the superlattice. The extra electrons (gathered in conduction valleys) and holes (gathered near peaks in valence band) are implied by the separation of chemical potentials, ϕ_n and ϕ_p , for the two species. Recombination is slowed by the spatial separation of electrons from holes. After ref. 1.

been produced by growing a crystal of GaAs by the technique of molecular beam epitaxy (MBE). Planar layers of ~ 40 nm thickness (~ 150 atomic layers) were doped alternately with silicon (n-type) and beryllium (p-type) atoms, at densities of about 10^{18} atoms per cm^3 . The resulting structure of the superlattice crystal is shown (Fig. 2) in the excited state resulting from illumination, which puts extra holes in the valence band and extra electrons in the conduction band. The extra electrons and holes are spatially separated with the interesting consequence that they can have anomalously long lifetimes, their mutual annihilation or recombination is slowed by their inability to approach each other. A second consequence is that some of the bending of the bands is undone by the extra carriers, leading to a crystal with a 'bandgap' uniquely sensitive to illumination.

Döhler *et al.* have produced unambiguous evidence for this sensitivity (Fig. 3). The quantum energy of the photoluminescence measures this bandgap, since the luminescence is annihilation radiation. At low excitation intensity, the bands shown in Fig. 2 are strongly bent and the gap energy is small, around 1.3 eV; but as the excitation intensity is increased, the gap energy shifts higher, approaching the energy of the bandgap of undoped GaAs, at about 1.53 eV.

Further evidence for the formation of a



1. Döhler, G.H. *et al.* *Phys. Rev. Lett.* **47**, 864 (1981).
2. Döhler, G.H. *Phys. Status Solidi* **52**, 79, 533 (1971); *J. Vac. Sci. Technol.* **16**, 851 (1979).
3. Esaki, L. & Tsu, R. *IBM J. Res. Devel.* **14**, 61 (1970).
4. Romanov, Yu. A. *Sov. Phys. Semiconductors* **5**, 1256 (1972); Romanov, Yu. A. & Orlov, L.K. *Sov. Phys. Semiconductors* **7**, 182 (1973).
5. Ludeke, R., Esaki, L. & Chang, L.L. *Appl. Phys. Lett.* **24**, 417 (1974).
6. Dingle, R., Wiegmann, W. & Henry, C.H. *Phys. Rev. Lett.* **33**, 827 (1974).
7. For a recent review, see Gossard in *Thin Films: Preparation and Properties* (eds Tu, K.N. & Rosenberg, R.) (Academic, New York, 1981).
8. Sai-Halasz, G.A., Tsu, R. & Esaki, L. *Appl. Phys. Lett.* **30**, 651 (1977).
9. Dingle, R. *Adv. Solid State Phys.* **15**, 21 (1975).
10. Klitzing, K.V., Dorda, G. & Pepper, M. *Phys. Rev. Lett.* **45**, 494 (1980).
11. Tsui, D.C., Gossard, A.C., Feld, B.F., Cage, M.E. & Dziuba, R.F. *Phys. Rev. Lett.* **48**, 3 (1982).

high-quality doping superlattice is given in the form of inelastic light-scattering spectra, which show how the electrons are distributed in their valleys, and how these distributions can be changed by optical excitation.

The techniques used in growing semiconductor superlattices are potentially of great importance. MBE, which originated at Bell Laboratories in the late 1960s, is one of a class of epitaxial crystal growth processes in which new material forms in thin layers on top of and in register with an existing crystalline substrate. These processes differ from each other in the manner in which new atoms are brought to the growth surface. Although other epitaxial processes have grown excellent crystals, precise control down to monatomic layers is still attained only with the MBE technique. Of course, it is not only superlattices which are grown: many of the most important structures involve simply one or two interfaces or heterojunctions between dissimilar semiconductors.

Among the important optical devices that have been fabricated starting with epitaxial growth are light-emitting diodes (LEDs), lasers and detectors; all of these are useful in the expanding field of optical communications, while LEDs also have important applications as display devices.

Because of the precise control which is possible in MBE growth and because of the invention, at Bell Laboratories, of modulation doping, researchers have been able to construct extremely smooth planar (two-dimensional) pathways for electronic conduction. Highly mobile electrons travelling in these layers may be involved in tomorrow's high-frequency electron devices.

But equally important is the 'pure physics' excitement which has attended the refinement of the MBE technique. Some of the earliest experiments were used to demonstrate the quantum-mechanical particle-in-a-box behaviour of electrons and Wannier excitons in AlGaAs-GaAs superlattices⁹. More recently some

unexpected peculiarities of matter in two-dimensional space (2D) have received attention. The effects of randomness or roughness of the pathway have a powerful influence on the localization of mobile particles in 2D, and studies of electron localization in epitaxial layers have increased our understanding of this concept.

A startling effect in 2D transport was discovered in 1980 by Klaus von Klitzing and his co-workers¹⁰. Their anomalous

Hall effect has been developed in MBE-grown AlGaAs-GaAs heterostructures to give a new precision to the value of the fine structure constant¹¹, and has also led to deeper theoretical insight into matter of reduced dimensionality.

Many more examples of exciting developments could be adduced, showing that new structures can be exploited both for new devices and for new science. I expect the doping superlattice to follow in this honourable tradition. □

Isotopic anomalies in meteorites

from Edward R.D. Scott

WHAT D.D. Clayton¹ has called a new field of astronomy, 'astrochemistry', is developing from an appreciation² that the Solar System did not form from a hot well-mixed cloud of uniform chemical and isotopic composition, as was once believed. Exciting discoveries of isotopic anomalies in meteorites, made during the past decade, are being used to test and develop theories for the formation of elements in stars, the formation and history of interstellar dust and the origin of the Solar System. Studies of the isotopic composition of oxygen and magnesium in chondrites provide crucial evidence for a dust component in the early Solar System containing essentially pure ¹⁶O, and ²⁶Mg excesses that result from the decay of ²⁶Al. Relatively simple, attractive and testable models have been proposed to explain the oxygen isotopic abundances, but the interpretation of the ²⁶Mg excesses is more controversial.

R.N. Clayton and his associates³⁻⁶ have established that the oxygen isotopic compositions of minerals and components in chondrites are a result of various degrees of mixing between reservoirs with intrinsically different isotopic compositions. On a plot of ¹⁷O/¹⁶O against ¹⁸O/¹⁶O, data for chondrules and nearly all Ca-Al-rich inclusions define a set of closely spaced lines with slopes of 0.94-0.97. (Terrestrial samples define a single line with a slope of 0.52, because of mass-dependent fractionation processes.) Three lines have so far been identified for H-L-L, CM and CO-CV chondrite groups, but others may be resolved. On the three-isotope plot, the lines appear to radiate from a point marking the oxygen composition of spinels, which occur in Ca-Al-rich inclusions and have the lowest ¹⁷O/¹⁶O and ¹⁸O/¹⁶O ratios of all minerals. Isotopic compositions of minerals may also depend on the nature of the host chondrule or inclusion; the largest variations are found

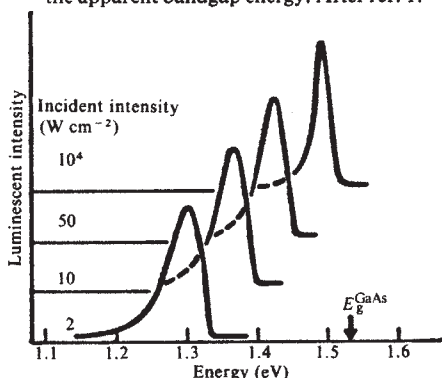
in the Ca-Al-rich inclusions of carbonaceous chondrites.

Oxygen isotopic data suggest that meteorite components largely formed from two ingredients: interstellar gas with a slightly variable isotopic composition, and well-mixed interstellar dust, about 5 per cent of which contained essentially pure ¹⁶O (refs 3,7,8). Dust, or chondrules and inclusions derived from this dust, exchanged oxygen with the gas to various degrees. Minerals such as spinel are presumed to preserve the original isotopic composition of the dust because of their low oxygen diffusion rates. (Oxygen diffusion rates in these minerals have not yet been measured.) FUN inclusions, which are those with many large isotopic anomalies², may have formed from poorly mixed interstellar dust balls⁸. During their formation, these inclusions experienced large isotopic fractionation, before partial exchange of oxygen with the gas⁹. Oxygen shows the most widespread isotopic anomalies because it was the only element that was abundant in both nebula gas and dust⁸.

Wood⁸ deduces from these and other data that the Ca-Al-rich inclusions and chondrules did not form by condensation from a gaseous solar nebula¹⁰, as many once believed. He supports D.D. Clayton's view¹¹ that they formed from unvaporized aggregates of interstellar grains. Ca-Al-rich inclusions probably formed either as distillation residues during unspecified heating events in the solar nebula⁸, or as condensates from supernova ejecta¹¹.

In contrast, the picture obtained from studies of magnesium isotopic abundances is relatively confusing. Excesses of ²⁶Mg formed by decay of ²⁶Al have been found in several minerals with high Al/Mg ratios¹². About 15 Ca-Al-rich inclusions in the Allende meteorite have been found to contain ²⁶Mg excesses that are correlated with the Al/Mg ratio. However, the proportion of their Al atoms that were radioactive when the inclusions formed (the initial ²⁶Al/²⁷Al ratio) varied widely from $< 2 \times 10^{-7}$ to 10^{-3} (see ref. 12). Some types of inclusion (type B1) have values of

Fig. 3 Optical spectra of the recombination luminescence emitted from a doping superlattice, showing how the increasing incident intensity shifts the radiation to higher frequency, implying that larger populations of electrons and holes change the apparent bandgap energy. After ref. 1.



Edward R.D. Scott is at the Department of Geology and Institute of Meteoritics, University of New Mexico, Albuquerque, New Mexico 87131.

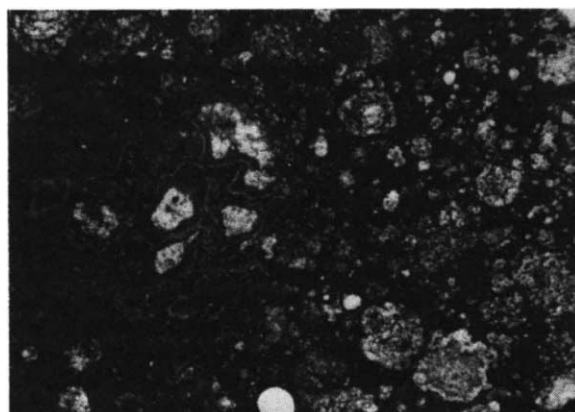
5×10^{-5} for this ratio, but in other inclusions, various crystals or minerals may give diverse values¹³. These differences may be due to a difference of up to 9 Myr in their times of formation (or final internal equilibration). Alternatively, $^{26}\text{Al}/^{27}\text{Al}$ ratios may have varied widely in the solar nebula¹². However, the size of isotopic variations that would be needed to account for simultaneous formation of these inclusions is much larger than has been found for other elements.

Bar-Matthews *et al.*¹⁴ have recently analysed an unusual Ca-Al-rich inclusion, which contains corundum (Al_2O_3), in the Murchison meteorite. If the inclusion had formed by equilibrium condensation from a slowly cooling gas of solar composition (which now seems rather unlikely), corundum should contain a large ^{26}Mg excess, as it is the first major mineral predicted to condense¹⁰ and has a very high Al/Mg ratio. Instead, the authors found only small ^{26}Mg excesses, which were uncorrelated with the Al/Mg ratio and which they attributed to Mg isotopic heterogeneities.

The lack of excess ^{26}Mg in Murchison corundum and the wide variation in the initial $^{26}\text{Al}/^{27}\text{Al}$ ratios of Allende inclusions are cited by D.D. Clayton^{1,15} as evidence against decay of ^{26}Al in the Solar System. He rejects explanations of time differences or isotopic heterogeneities and argues that ^{26}Al decayed in interstellar grains. He attributes correlated ^{26}Mg excesses in individual anorthite crystals to precursor Al_2O_3 stardust. This controversial idea has stimulated more precise experiments, but is still a minority view.

Analogous problems have arisen in the interpretation of ^{129}Xe excesses in ordinary chondrites due to decay of ^{129}I , which has a half life of 17 Myr. If ^{129}I was homogeneously distributed in the solar nebula, then two chondrules in the Bjurböle chondrite formed 2 Myr apart¹⁶, a

Transmitted light photograph of the Vigarano carbonaceous chondrite showing Ca-Al-rich inclusion with complex rims and smaller chondrules. Actual size: 8×11 mm.



surprisingly long time. D.D. Clayton has also questioned the existence of live ^{129}I in the Solar System¹, but the best working hypothesis seems to be that the I-Xe time differences, like those for the Al-Mg chronometer, may be real. Some authors have invoked planetary, rather than nebular, processes to form chondrules^{17,18} and Ca-Al-rich inclusions¹⁹. This might help to explain the large spread of formation ages of chondrules and inclusions, but the oxygen isotopic data strongly suggest that chondrules and Ca-Al-rich inclusions are all primitive objects, which may have formed during similar

high-temperature events²⁰.

Progress in astrochemistry will be greatly advanced when the origin of the chondrules and Ca-Al-rich inclusions is better understood. This probably requires a realization that all the inhabitants in the exotic menagerie of chondritic components — Ca-Al-rich inclusions and chondrules, mafic chondrules, aggregates and matrix, and metallic Fe, Ni — have interrelated origins. The origin of chondrules in enstatite chondrites, for example, cannot be divorced from an understanding of the formation of Ca-Al-rich inclusions in carbonaceous chondrites.

A new model for nitrogen control

from M.J. Merrick

In *Escherichia coli* and other enteric bacteria, the induction of a number of operons involved in carbon assimilation, for example, *lac*, *ara*, *gal*, is affected by the intracellular concentration of cyclic AMP, which is in turn determined by the carbon source available to the cell. This phenomenon is termed catabolite repression and is mediated by a complex of cyclic AMP and the cyclic AMP receptor protein (CRP) which together form a transcriptional activator. There has long been evidence¹ for another regulatory system controlling the balance of nitrogen assimilation enzymes in bacteria, for example, histidine (*hut*) and proline (*put*) utilization, and nitrogen fixation (*nif*). This phenomenon, which has been termed 'nitrogen control'², has received considerably less attention than catabolite repression although a model for control of nitrogen assimilation was proposed in 1973 (refs 3,4). Recent evidence from genetic studies in a number of organisms suggests that the system is considerably more complex than originally proposed and involves both positive and negative regulatory elements.

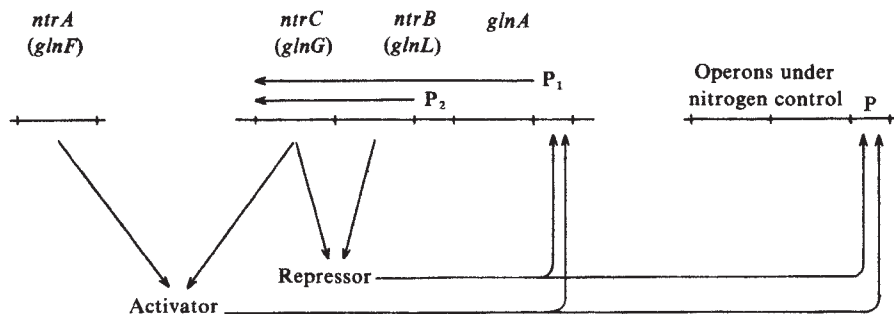
Studies on the regulation of histidase synthesis in *Klebsiella aerogenes* originally led Magasanik and co-workers to formulate a model⁵ in which the principal regulator of nitrogen assimilation was the

enzyme glutamine synthetase (GS). They postulated that GS was responsible not only for the biosynthesis of glutamine and the assimilation of nitrogen into glutamate, but also for the activation of transcription of operons such as *hut* and *put*. The enzymatic activity of GS is regulated by the reversible adenylation of specific tyrosyl residues on each subunit of this dodecameric protein. In conditions of nitrogen starvation, GS is deadenylylated and highly active, and Magasanik *et al.* proposed that this deadenylylated form of GS was an activator of transcription. Evidence for the model included the isolation of mutants, believed to be in the GS structural gene (*glnA*), which showed altered regulatory properties⁵.

In the last five years, studies of nitrogen control have been extended to *Salmonella typhimurium*, *E. coli* and *Klebsiella pneumoniae* as well as *K. aerogenes*. A very different model for regulation of nitrogen assimilation has now emerged in which GS is no longer considered to be a regulatory protein and instead three specific genes, *ntrA*, *ntrB* and *ntrC* (or their equivalents), have been identified in all the above four species. Two of these genes, *ntrB* and *ntrC*,

M.J. Merrick is in the Agricultural Research Council Unit of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ.

1. Clayton, D.D. *Q. Jl R. astr. Soc.* (in the press).
2. Begemann, F. *Rep. Prog. Phys.* **43**, 1309-1359 (1980).
3. Clayton, R.N., Onuma, N., Grossman, L. & Mayeda, T.K. *Earth planet. Sci. Lett.* **34**, 209-224 (1977).
4. Gooding, J.L. *et al. Meteoritics* **15**, 295 (1980).
5. Clayton, R.N., Mayeda, T.K., Gooding, J.L., Keil, K. & Olsen E.J. *Lunar planet. Sci.* **12**, 154-156 (1981).
6. Clayton, R.N. & Mayeda, T.K. *Lunar planet. Sci.* **13**, 117-118 (1982).
7. Blander, M. & Fuchs, L.H. *Geochim. cosmochim. Acta* **39**, 1605-1620 (1975).
8. Wood, J.A. *Earth planet. Sci. Lett.* **56**, 32-44 (1981).
9. Clayton, R.N. & Mayeda, T.K. *Geophys. Res. Lett.* **4**, 295-298 (1977).
10. Grossman, L. *A. Rev. Earth planet. Sci.* **8**, 559-608 (1980).
11. Clayton, D.D. *Astrophys. J.* **199**, 765-769 (1975); *Moon Planets* **19**, 107-137 (1978).
12. Lee, T. *Rev. Geophys. Space Phys.* **17**, 1591-1611 (1979).
13. Hutcheon, I.D. *Am. chem. Soc. Symp. Ser.* **176**, 95-128 (1982).
14. Bar-Matthews, M., Hutcheon, I.D., MacPherson, G.J. & Grossman, L. *Geochim. cosmochim. Acta* **46**, 31-41 (1982).
15. Clayton, D.D. *Lunar planet. Sci.* **13**, 115-116 (1982).
16. Caffee, M.W., Hohenberg, C.M., Hudson, B. & Swindle, T.D. *Lunar planet. Sci.* **13**, 75-76 (1982).
17. Leitch, C.A. & Smith, J.V. *Nature* **290**, 578-579 (1980).
18. Zook, H. *Lunar planet. Sci.* **12**, 1242-1244 (1981).
19. Armstrong, J.T., Meeker, G.P., Huneke, J.C. & Wasserburg, G.J. *Geochim. cosmochim. Acta* **46**, 575-595 (1982).
20. Scott, E.R.D., Taylor, G.J. & Keil, K. *Lunar planet. Sci.* **13**, 704-705 (1982).



Model for nitrogen control in *S. typhimurium* and *E. coli*.

are contiguous with the GS structural gene, *glnA*, and this close linkage confounded the earlier analysis of mutants.

The *ntrA* gene (originally designated *glnF*) was first identified in *S. typhimurium*⁶. *NtrA* is not linked to *glnA* but mutations in this gene result in constitutive low levels of GS and glutamine auxotrophy⁶⁻⁸. These mutants are also unable to activate other genes under nitrogen control.

Following identification of *ntrA*, another regulatory locus closely linked to *glnA* was described in *S. typhimurium*², *E. coli*⁹ and *K. pneumoniae*^{8,10,11}. Then at the 1981 meeting of the American Society for Microbiology in Dallas, it was reported that in both *S. typhimurium* (Kustu and colleagues at the University of California, Davis) and *E. coli* (Tyler and colleagues at the Massachusetts Institute of Technology), this regulatory locus consists of two cistrons (*ntrB* and *ntrC* in *S. typhimurium*, and *glnL* and *glnG* in *E. coli*). Tyler *et al.* have shown that in *E. coli*, the two genes constitute a single operon transcribed from *glnL* to *glnG* and it has been proposed that this operon may be transcribed either from its own weak promoter (P₂) or by stronger read-through transcription from the *glnA* promoter (P₁)¹². The products of the *E. coli ntrB* and *ntrC* genes have been identified as polypeptides of molecular weights 36,000 (*ntrB*)¹⁴ and 54,000 (*ntrC*)^{13,14}.

The glutamine auxotrophy induced by *ntrA* mutations can be suppressed by mutations in either *ntrB* or *ntrC* with a resultant constitutive low level of GS synthesis. Hence it seems that in *ntrA* mutants, GS is permanently repressed and this repression is relieved by *ntrB* or *ntrC* mutations, leaving a low level of unregulated *glnA* transcription from P₁. This suppression does not, however, extend to other operons under nitrogen control which fail to be activated in all *ntrA*⁻ strains. In an *ntrA*⁺ background, *ntrB* mutations result in constitutive expression of *glnA* whereas most *ntrC* mutations prevent expression of *glnA* and other genes under nitrogen control.

From the extensive genetic analysis in both *S. typhimurium* and *E. coli* a consensus model (see the figure) has emerged for nitrogen control in the two organisms. It is proposed that the *ntrB* and *ntrC* products can repress or activate trans-

cription of genes under nitrogen control, including *glnA*. Repression requires both the *ntrB* and *ntrC* products (possibly acting as a complex), hence the suppression of the *ntrA* Gln phenotype by mutations in either *ntrB* or *ntrC*. However, activation apparently requires only the *ntrC* (and *ntrA*) product(s) because *ntrB* mutants express *glnA* constitutively. The Gln phenotype of certain *ntrC* mutants is proposed to be due to the *ntrB*-*ntrC* complex being permanently locked in the repressor form, a model which is supported by the isolation of Gln⁺ suppressors of such *ntrC* mutations in both *ntrB* and *ntrC*¹⁵.

The *ntrC* product therefore seems likely to be a DNA-binding protein which can activate transcription. The *ntrB* product could also bind to DNA in the manner of a conventional repressor or alternatively could modify the activator properties of the *ntrC* product. The *ntrA* product is a positive regulatory factor necessary for the formation of a functional *ntrC* activator and hence *ntrA* mutants cannot derepress genes under nitrogen control. Analogies between this system and the cyclic AMP-CRP system for carbon and energy regulation are striking and extension of such analogies suggests that *ntrA* may encode an enzyme that synthesizes a low molecular weight co-regulator of transcription — a signal of nitrogen deficiency². It is of interest that certain operons, for example, *hut*, are subject to regulation by catabolite repression and by nitrogen control mediated by the elements described above.

The *ntrB*, *ntrC* regulatory system in which two genes in a single operon encode a repressor and an activator respectively is almost unique among the bacterial control systems so far analysed. However, one other system shows striking analogies to that proposed for *ntrB* and *ntrC*.

The nitrogen fixation (*nif*) regulon in *K. pneumoniae* is controlled by two *nif*-specific regulatory genes, *nifL* and *nifA*, which act respectively to repress or activate *nif* transcription^{16,17}. These two genes comprise a single operon, *nifLA*, and their products may also function as a protein complex. As with *ntrB,C*, the product of one *nif* gene, *nifA*, is alone sufficient for activation of transcription¹⁸. Interestingly the *nifLA* operon is in turn regulated by the general nitrogen control system which our recent studies have shown to be essentially

identical to that described for *ntrA,B,C* in *E. coli* and *S. typhimurium*.

The recent work described above should stimulate new interest in the regulation of nitrogen assimilation, but just as the mechanism of catabolite repression is still not completely understood¹⁹, the details of nitrogen control will probably take considerable time to dissect. In the meantime, a recent report of a regulatory locus closely linked to the structural gene for GS in *Bacillus subtilis*²⁰ prompts speculation as to whether nitrogen control in organisms other than the enteric coliform bacteria is mediated by a similar genetic system. □

1. Neidhart, F.C. & Magasanik, B. *J. Bact.* **73**, 253 (1957).
2. Kustu, S., Burton, D., Garcia, E., McCarter, L. & McFarland, N. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4576 (1979).
3. Brenchley, J.E., Prival, M.J. & Magasanik, B. *J. biol. Chem.* **248**, 6122 (1973).
4. Prival, M.J., Brenchley, J.E. & Magasanik, B. *J. biol. Chem.* **248**, 4334 (1973).
5. Magasanik, B. *et al. Curr. Topics Cell Regulation* **8**, 119 (1974).
6. Garcia, F., Bancroft, S., Rhee, S.G. & Kustu, S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1662 (1977).
7. Gaillardin, C.M. & Magasanik, B. *J. Bact.* **133**, 1329 (1978).
8. Leonardo, J.M. & Goldberg, R.B. *J. Bact.* **142**, 99 (1980).
9. Pahel, G. & Tyler, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4544 (1979).
10. de Bruijn, F.J. & Ausubel, F.M. *Molec. gen. Genet.* **183**, 289 (1981).
11. Espin, G., Alvarez-Morales, A. & Merrick, M. *Molec. gen. Genet.* **184**, 213 (1981).
12. Magasanik, B. & Rothstein, D.M. in *Glutamine: Metabolism, Enzymology and Regulation* (eds Mora, J. & Palacios, R.) 61 (Academic, New York, 1980).
13. Backman, K., Chen, Y.-M. & Magasanik, B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3743 (1981).
14. McFarland, N., McCarter, L., Artz, S. & Kustu, S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2135 (1981).
15. Wei, G.R. & Kustu, S. *Molec. gen. Genet.* **183**, 392 (1981).
16. Dixon, R. *et al. Nature* **286**, 128 (1980).
17. Merrick, M. *et al. Molec. gen. Genet.* (in the press).
18. Buchanan-Wollaston, V., Cannon, M.C., Beynon, J. & Cannon, F.C. *Nature* **294**, 776 (1981).
19. Ullmann, A. & Danchin, A. *Trends biochem. Sci.* **5**, 95 (1980).
20. Reysset, G. *J. Bact.* **148**, 653 (1981).



The Photographic Gun — It takes twelve images, in one second, of an object on which the piece is continuously sited. The exposure time of each image is 1-720th of a second. It is being applied to the taking of birds flying.

From *Nature* **26**, 25 May 1882.

Vitamin D: sunlight and precursors

from Alastair Hay

WITH some twenty metabolites of vitamin D now identified (see the figure for some of them) chemists in the field have much to be satisfied about. There is less satisfaction, however, for those concerned with the biological importance of these metabolites and a recent workshop* on vitamin D raised more questions than it answered.

Take, for example, the role of the two kidney metabolites of vitamin D, 1,25-dihydroxyvitamin D ($1,25-(OH)_2D$) and 24,25-dihydroxyvitamin D ($24,25-(OH)_2D$). The former is the active hormonal form of vitamin D. The controversial suggestion that $24,25-(OH)_2D$ is also a potential hormone active in bone was made several years ago but it is still far from clear whether $1,25-(OH)_2D$ alone is necessary to cure the bone disease osteomalacia — caused by a deficiency of vitamin D — or whether both metabolites are required.

Obtaining an adequate supply of vitamin D is a greater problem for old people than for the young. The two main sources are from food and exposure to UV light in the summer. For most people in Britain, including the elderly, it is the latter source which is the most important (Lawson *et al. Br. med. J.* II, 303; 1979).

Although both young and old are said to produce equivalent amounts of vitamin D when exposed to UV light (Davie & Lawson *Clin. Sci.* 58, 235; 1980), the concentration of the precursors of vitamin D₃ in the skin seems to fall with age. M.F. Holick (Massachusetts General Hospital) reported that the epidermal stores of 7-dehydrocholesterol (7-DHC) are inversely related to age, as is the amount of pre-vitamin D₃ (pre-D₃) formed from 7-DHC after UV irradiation.

One possible regulator of 7-DHC stores is the kidney metabolite $1,25-(OH)_2D$. Holick has detected a high-affinity, low-capacity receptor for $1,25-(OH)_2D$ in fibroblasts and keratinocytes cultured from skin biopsies taken from normal volunteers that was absent in cells cultured from a skin biopsy of a patient with vitamin D-dependent rickets type II. The kidney metabolite increased the ratio of 7-DHC to cholesterol in cell cultures of keratinocytes and Holick suggested that stores of 7-DHC in skin may be controlled by $1,25-(OH)_2D$, inhibiting the conversion of 7-DHC to cholesterol.

The steps in the conversion of 7-DHC to pre-D₃ and then vitamin D₃ are controlled by UV light and temperature respectively. The vitamin is removed from the skin attached to a specific vitamin D-binding

protein (DBP) which has a much higher affinity for vitamin D₃ than for pre-D₃ and so allows only the former to enter the circulation.

There has been much speculation about the function of DBP and argument about the importance of free and bound vitamin D₃ metabolites in plasma. It now seems clear that the major role of DBP is as a storage protein for vitamin D metabolites, for which it has a high affinity.

Recently Haddad *et al. (J. clin. Invest.* 67, 1550; 1981) suggested that vitamin D and its metabolites attached to DBP might first bind to a membrane receptor and then be pinocytosed into cells. R. Bouillon (University of Leuven) reported, however, that he could find no sign of a membrane receptor for DBP. Labelled DBP did not bind to kidney cells in culture in any appreciable quantity and the little that did bind was not altered by the presence of either vitamin D metabolites or unlabelled DBP. The binding was therefore not specific, a finding which Bouillon confirmed using immunofluorescent techniques.

Thus Bouillon concluded that DBP acts like other transport proteins for hormones and that it is the free hormone which crosses cell membranes. In normal subjects, a positive correlation was found between the serum concentrations of DBP and $1,25-(OH)_2D$. As vitamin D status does not affect DBP levels, the correlation suggests that $1,25-(OH)_2D$ levels are related to changes in DBP concentration and that the free, rather than the total, $1,25-(OH)_2D$ concentration is feedback regulated.

Despite a 1,000-fold difference between the total concentrations of the liver metabolite 25-hydroxyvitamin D and $1,25-(OH)_2D$, the difference between the free concentrations and these metabolites is only 10-fold (Bouillon). If it is the free hormone which is important, it is clear that some thought will have to be given to the value of the current assays which measure total concentrations of vitamin D

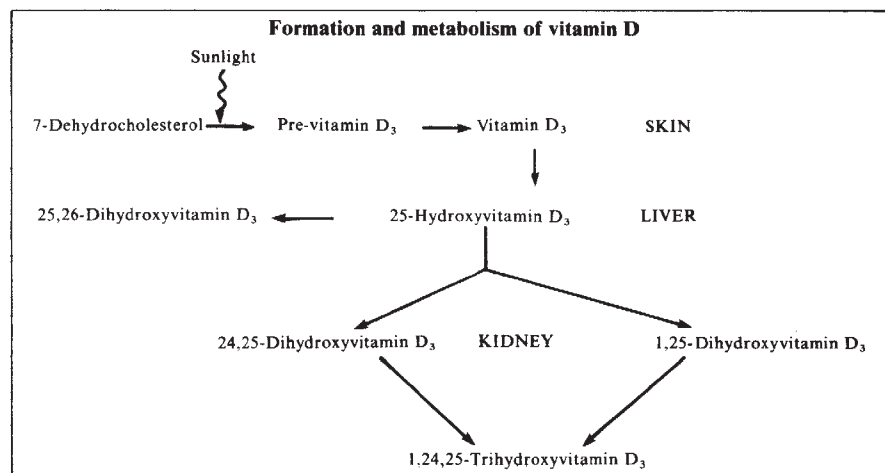
metabolites.

Another function of $1,25-(OH)_2D$ which is becoming increasingly controversial is the hormone's action on the gut, where it increases the active transport of calcium. R.T. Franceschi and H.F. De Luca (*J. biol. Chem.* 256, 3848; 1981) and D.E.M. Lawson (Dunn Nutrition Laboratory, Cambridge) believe that it is new proteins synthesized in the mucosa in response to $1,25-(OH)_2D$ which facilitate calcium transport across the gut. But H. Rasmussen (Yale University) argues that the increase in calcium transport is not affected when protein synthesis is inhibited. He has observed that $1,25-(OH)_2D$ increases the synthesis of phosphatidylcholine in mucosal cells — in the absence of protein synthesis — and he believes that this change is necessary for active calcium transport.

Rasmussen claims that the increased phosphatidylcholine concentration is eventually reflected in changes in the phospholipid composition of the brush border membrane. D.D. Bikle (University of California) would probably support this view. He ruled out newly synthesized protein within brush border membrane, or altered phosphorylation of proteins in response to $1,25-(OH)_2D$, as the explanation for the increased calcium flux across the membrane. He found that calcium transport occurred in the absence of both these changes.

J.A. Putkey (University of California), however, could find no evidence to support the lipid hypothesis. Use of a spin label probe, 5-nitroxide stearate, showed that dietary vitamin D has no detectable effect on either the polarity, or lipid fluidity, of the membrane; neither does it alter its cholesterol or lipid phosphorus content. It seems, however, that some consensus view, which relies on both protein synthesis and lipid turnover to explain the action of $1,25-(OH)_2D$ in the gut, is sure to emerge sooner or later.

If just one or two of vitamin D's metabolites can generate so much controversy, then the chemists would seem to have provided ample ammunition to allow the disputes to continue for some time yet. □



*The 5th Workshop on Vitamin D was held at Historic Williamsburg, Virginia on 14–19 February 1982. Proceedings will be published by Walter de Gruyter later this year.

REVIEW ARTICLE

Variety in the level of gene control in eukaryotic cells

James E. Darnell Jr

The Rockefeller University, New York, New York 10021, USA

In light of present progress in the understanding of how messenger RNA is constructed in eukaryotic cells, the levels of gene control are discussed. Transcriptional control, assessed by the rate of synthesis of specific nuclear RNA, is strongly indicated to be the most frequent level of control. Definition of eukaryotic transcription units as simple (encoding one protein) and complex (encoding two or more proteins) affords a framework in which to discuss control at the level of RNA processing for which several examples also are known. Finally, differential stability of cytoplasmic mRNAs and differential translation, both well established, are briefly described.

THE study of gene control in eukaryotic cells has matured over the past five years into one of the most active and productive fields in experimental biology. Earlier conclusions about eukaryotic gene control were deferred because mRNA production in eukaryotes is not a simple matter of transcribing the RNA. Especially in vertebrate cells, the transcription units for mRNAs are longer than the final mRNA product¹⁻⁶. Several 'processing' steps occur to primary RNA transcripts in the cell nucleus before the mRNA emerges to be translated in the cytoplasm, involving the addition of a methylated guanylate residue—a 'cap'—at the 5' end of primary transcripts⁷, the addition of adenylic acid residues to a site in the primary transcript to create a 3' poly(A) segment⁸⁻¹³ and finally the removal of specified internal sequences (introns)¹⁴ and rejoining or splicing of the remaining RNA pieces (exons)¹⁴ to create a finished mRNA (refs 15-17; Fig. 1).

This complicated mRNA formation, together with the ordered, complex structure of the nucleus and cytoplasm of eukaryotic cells and the different fundamental needs served by a genetic programme in slowly growing or non-growing differentiated eukaryotic cells and rapidly adaptable bacterial cells, have long spurred speculation^{5,18,19} that eukaryotic gene control might differ from that in bacteria. In particular, it has been suggested that eukaryotic genes are subject to regulation at more levels than are bacterial genes. In addition to transcriptional regulation, there might be regulation of RNA processing in the nucleus, mRNA transport from the nucleus, and mRNA stability and frequency of translation in the cytoplasm.

With the major mechanical steps of eukaryotic mRNA formation now understood, it is possible to describe more accurately the anticipated levels and varieties of eukaryotic gene control⁵. In fact, regulation of the amount of specific mRNA molecules has been demonstrated at the level of transcription, nuclear RNA processing and mRNA stability. Here I shall cite evidence supporting a variety of levels of eukaryotic gene control. The experiments reviewed mainly demonstrate the levels of control mRNA concentration, and not the biochemical mechanisms underlying the controls. No attempt is made to review the control of transcription units encoding ribosomal RNA, transfer RNA or 5S RNA (RNA polymerase I and III transcription units; transcription of 5S RNA has been recently reviewed²⁰). The relationship between gene control and chromatin structure is not discussed but has been the subject of a review recently published²¹. It is clearly possible to envisage gene control at the level of differential translation of mRNAs and potential differential post-translational events, but these subjects have received less experimental attention, so they will be mentioned only briefly.

RNA transcription and transcriptional control

The first level of control in eukaryotic cells is in the choice of RNA polymerases. In contrast to bacteria, eukaryotes have three RNA polymerases: polymerase II transcribes genes that encode proteins^{22,23}, polymerases I and III have already been mentioned. Both *in vivo*²⁴ and *in vitro*^{23,25}, RNA polymerase II initiates transcription at a site located at least in part by a conserved 8-10 nucleotide region 25-30 nucleotides upstream (in the 5' direction) from the RNA start site. The nucleotides TATA are strongly conserved at this site^{23,26} and the region is known as a 'TATA' or Goldberg-Hogness box²⁷. Undoubtedly other sequences are important in regulating access of polymerases to DNA. Reports on viral, yeast and individual mammalian genes point to the importance of sequences upstream

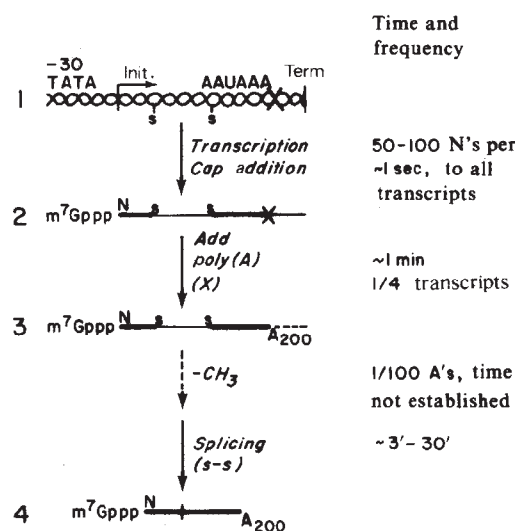


Fig. 1 Steps in mRNA formation. (1) DNA indicating the sites of initiation (INIT.), about 30 nucleotides to the 3' site of a TATA box, and of termination (Term.) of transcription. (2) Initiated transcript with cap (m^7Gppp) added to the first nucleotide (N). (3) Transcript after polyadenylation—the addition of a sequence of about 200 adenylic residues (A_{200}) at the poly(A) site, which is 10-20 nucleotide bases downstream from the conserved hexanucleotide AAUAAA. Between steps (2) and (4), mammalian mRNAs acquire methyl groups ($-CH_3$) on the 6 position of about 1 of every 100 adenylic residues. The exact time of methyl group addition is not known. (4) Messenger RNA formed by splicing. X, poly(A) site; s, splice site.

from the TATA box (see, for example, refs 28–33). Although no consistent picture has yet emerged for the role of these upstream sequences in transcriptional control, intense research in this area should yield important conclusions soon.

Here we are concerned more with establishing the variation of transcription rates than with the basis for the variation. Transcription initiation is widely believed to be the major controlling step in transcriptional control. Little is known, however, about transcriptional termination, which might also be a level of control. For example, all RNA polymerase II complexes that begin a chain do not transcribe the entire chain. Prematurely terminated transcripts are produced from viral³⁴ and probably cell³⁵ transcription units. Little is known about the specificity of transcriptional termination sites although polymerases seem regularly to pass poly(A) sites^{10–12}. Recent evidence suggests that, for the major late adenovirus^{10,11} and the major β -globin transcription units¹⁵⁰ there exists a region of termination, if not a specific site.

Transcriptional control

The various mechanisms for transcriptional control are very familiar from the classic studies of bacterial gene regulation^{36,37}. In bacteria, both regulated increases and decreases of transcriptional initiation are possible and two types of regulation of termination can occur—'readthrough' to the next gene and premature termination or 'attenuation' which stops transcription before the coding portion of the gene is reached³⁸. Apparent control of transcriptional initiation that results in an increase or decrease of mRNA has been documented in several cases in eukaryotic cells and one case of differential termination has been described. The evidence reviewed here shows clearly that the most frequent mode of eukaryotic gene control is at the transcriptional level.

Measurement of transcription control. Bacterial geneticists used RNA–DNA hybridization to document the existence and very short half life of mRNA³⁹ and to map transcription units^{37,38,40,41}. The simple presence or absence of specific mRNAs was taken as evidence (correctly in most cases) for transcriptional control. However, in eukaryotic cells, a changing level in the steady-state concentration of a given mRNA or an increased labelling of mRNA after relatively long label times could partly be due to differential nuclear or cytoplasmic stabilization. To avoid this difficulty it is necessary to hybridize very briefly labelled ('nascent') nuclear RNA to DNA sequences corresponding to a specific mRNA⁵. The ideal method to label nuclear RNA is expose cells to radioactive RNA precursors for say ≤ 5 min, and thereby sample the labelled nuclear RNA even before the acid-soluble triphosphate pool has become fully labelled^{10,35}. But this is often not practicable because so little of the total RNA output (usually 5×10^{-6} – 5×10^{-5} of the total) from the nucleus is due to transcription of a single gene^{4,42,43}. Fortunately there is another suitable method for labelling nascent nuclear RNA. In isolated nuclei, previously initiated RNA polymerase II molecules elongate growing RNA chains by less than 500 nucleotides in 10–20 min of incubation^{1,4,44}. RNA synthesis in isolated nuclei depends on nucleoside triphosphates (NTPs) and thus nuclear RNA can be highly labelled with high-specific activity NTPs. Two studies suggest that *in vitro* chain elongation is an accurate reflection of a brief label inside cells. First, in the cases of adenovirus-2 transcription units¹ and the major β -globin transcription unit⁴, the polymerases in isolated nuclei elongate intact molecules covering the entire transcription unit. Second, the fraction of total *in vitro* or *in vivo* labelled RNA complementary to various specific DNA segments is similar^{1–4,45–48}.

Therefore isolated nuclei are widely used to prepare nascent RNA for the measurement of differential transcription of specific genes. All the experiments I cite as establishing transcriptional control have used RNA from either pulse-labelled cells or isolated nuclei to estimate transcriptional rate.

Adenovirus transcriptional control. The changes in virus protein and virus mRNA synthesis during adenovirus infection have

been thoroughly explored (Fig. 2). Specific increases and decreases in rates of transcription, presumably mediated at the initiation step, occur at several sites on the adenovirus genome during virus infection. In addition, one case of apparent termination control has been described. The following conclusions can be drawn about adenovirus transcriptional controls.

(1) During the first 1–2 h of infection, transcription units 1A, 1B, 3 and 4 all increase in transcription rate⁴⁹. Of this group, transcription of 1A appears to peak first. In comparison, transcription unit 2 is delayed^{49,50} and the transcription unit for protein IX produces no mRNA or nuclear RNA at all for the first 4–6 h (refs 51, 52), although the protein IX transcription unit resides entirely within the early transcription unit 1B (refs 26, 52, 53; Fig. 2). At ~6–8 h after infection, protein IX transcription increases without affecting the rate of transcription of 1A or 1B (ref. 52). Among the tests for early activity of the protein IX transcription unit was an assay for promoter-proximal RNA capable of detecting minimal (~1%) synthesis of the first 100 bases of the transcription unit, but still no activity was evident⁵². This is one of the clearest cases in eukaryotic cells or viruses of transcriptional control at the level of initiation.

(2) Mutations in transcription unit 1A lead to a failure during infection of mRNA accumulation from regions 1B, 2, 3 and 4 (refs 54–57). This failure is due to the absence of transcription when a 1A protein product is defective⁵⁷, but the exact mechanism of action of the 1A protein remains unknown⁵².

(3) P16 (the major late adenovirus promoter) functions both early^{59–62} and late^{1,3,24,34} in infection, but after DNA synthesis, it is 100–1,000 times more active per cell^{34,60}. This increased transcription is presumably on new DNA templates. In comparison, transcription units 1A and 1B on the same strand remain equally active throughout infection^{34,52,63}. Thus there seems to be a mechanism for restricting polymerase initiation to P16 on the new DNA.

(4) Controlled termination of transcription also occurs during adenovirus infection^{60–63}. The same P16 promoter produces RNA beginning at the same cap site both early and late in infection. From the late transcripts, five 3' co-terminal groups of mRNA are formed^{10,64–66}, but early in infection only the L1 poly(A) site is found in mRNA^{60,62}. Early transcription does not seem to pass map unit 60 and perhaps ends between 40 and 50 (refs 60–62), whereas late transcription continues to around map unit 98 (refs 10, 11). This is equivalent to read-through transcriptional control^{37,38}. As there is an enormous increase in transcription from P16 when DNA is replicated it seems likely that transcription of the same DNA molecule is

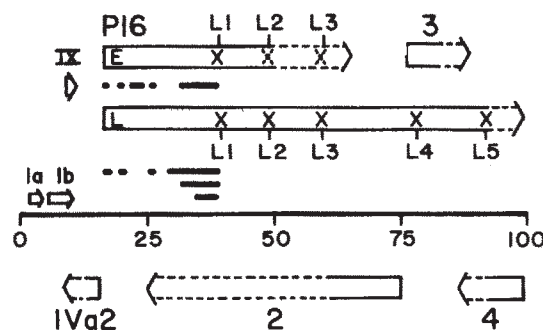


Fig. 2 Features of the transcription map of adenovirus transcription units (adapted from refs 5, 6, 52, 59–63). The genome is 36 kb long, so each map unit equals 360 bases. Each open bar, 1a, 1b, 1X, 3, represents a transcription unit. The transcription unit P16 beginning at map position 16 is shown both early (E) and late (L) in infection. The arrowheads show direction of transcription and dashed lines indicate cases where transcription proceeds beyond poly(A) sites. Only for P16 are the poly(A) sites marked (X). Special prominence is given to P16 to show that early in infection, transcription definitely terminates before reaching L3, L4 and L5; the thick black lines under P16E and P16L indicate the different regions retained after splicing in early and late L1 mRNAs.

not terminated in two different ways, but that one transcription unit in the adenovirus genome can have different termination sites early and late in infection.

(5) Transcription of region 4 peaks early in infection and then at 4–8 h declines to $\leq 10\%$ of the rate it reached by 4 h (ref. 67). Protein synthesis is necessary for this programmed decline, particularly a protein from region 2 as mutations in region 2 block the decrease.

SV40 transcriptional control. A specific decrease in the early transcription unit of simian virus 40 (SV40) has been documented during the later periods of infection⁶⁸. At the beginning of infection, the overall rate of transcription of SV40 is quite low and there is five times as much RNA complementary to the early transcription unit as to the late transcription unit. Late in infection, after a very large increase in total SV40 RNA in infected cells, there is only 1/20th as much RNA complementary to the early transcription unit as to the late transcription unit. This change in ratio of mRNAs from the two transcription units is probably due to the T antigen, the protein encoded by the early transcription unit⁶⁹. T antigen has been shown to bind three identical sites within a 100–150 base pair region that overlaps the early promoter^{70,71}. This binding is necessary for DNA replication⁷² but also stops or slows considerably the early transcription of SV40. The effect of T antigen is specific to the early promoter as transcription of recombinant DNA molecules, in which the adenovirus major late promoter has been inserted no more than 50 bases away from the early SV40 promoter, is unaffected by the *in vitro* binding of the T antigen to the SV40 promoter⁷¹.

Hormone-induced increases in transcription. Many proteins in differentiated cells are known to be produced in increased amounts in tissues stimulated by the appropriate hormone. Transcription in isolated nuclei shows that this stimulation is due, at least mainly, to transcriptional control.

One example is the stimulation of production of egg white proteins in the chicken oviduct by oestrogen. cDNA clones to the egg white proteins ovalbumin^{73,74}, conalbumin⁷⁴ and ovomucoid⁷⁵ all hybridize 10–50-fold more labelled nuclear RNA from the nuclei of hormone-treated chick oviduct epithelial cells than from control cells. Nuclei from other tissues such as liver or spleen do not make labelled RNA complementary to ovalbumin or ovomucoid cDNA. Maintenance of a high level of transcription in the oviduct requires continuous hormone treatment.

Liver cells of roosters also respond to oestrogen treatment by producing vitellogenin and very low density lipoprotein (VLDL)^{76,77}. The cDNAs corresponding to these mRNAs detect a greatly increased rate (>50 -fold) of RNA synthesis in chromatin fraction (lysed nuclei) from treated compared with untreated animals. Thus increased transcription of specific genes by steroid hormones seems well established.

In addition to these cases of cell genes, RNA synthesis from the integrated mammary tumour virus DNA also is increased by oestrogens²⁹.

Transcriptional control of chick globin. For the first few days after fertilization, chick embryos synthesize embryonic forms of α - and β -globin. Later on adult forms of the two globins are synthesized. The same cell does not differentiate first to form embryonic and then an adult globin. Rather, erythroblastic cells choose one of the two pathways^{78,79}. Genomic DNA clones have been used to show that isolated nuclei from 5- or 12-day-old embryos produced labelled RNA complementary only to embryonic or adult globin DNA respectively. Thus the differential expression of these two forms of chick globin appears to be controlled at the level of transcription.

Liver-specific transcriptional control. A series of cDNA clones prepared from mouse liver mRNA were screened to select mRNA sequences present in many cell types ('common' mRNAs) and mRNA sequences present only or mainly in liver⁴⁶. Nuclei from liver cells, cultured undifferentiated cells and brain cells were then used to prepare labelled nascent RNA. All 11 liver-specific mRNA sequences were labelled in

liver nuclei but not in brain or cultured cell nuclei. Sequences complementary to the common cDNA clones were labelled in all nuclei. As the clones used in the test represented mRNAs that ranged from ~ 100 to $>10,000$ copies per cell, it seems that, in general, specific mRNAs in liver cells are controlled transcriptionally. A further experiment with liver nuclei reveals a specific case of changing transcriptional controls. Adult male mice excrete a protein in urine, the major urinary protein (MUP), which is not excreted by newborn mice⁸⁰. Liver nuclei from adult and newborn animals were used to score transcription of MUP genes. The adult nuclei synthesized at least 100-fold more MUP-specific RNA than did the nuclei from the livers of newborns⁸¹.

Transcriptional control by choice of initiation sites. More of the same α -amylase enzyme protein is produced in rat salivary glands than in liver. When the cDNAs complementary to α -amylase mRNA from the two tissues were sequenced, the 5' nucleotides of the two mRNAs were found to differ while the remainder of the mRNA chains were identical⁸². Examination of genomic DNA⁸³ and nuclear RNA (U. Schibler, personal communication) revealed that the cap sites for the two amylase mRNAs lie ~ 2.8 kilobases apart but the two overlapping transcription units cover the same protein coding region and the same poly(A) site(s). This necessitates a different splice to join the two different cap and leader sequences to the same mRNA 'body'. However, as cap sites always seem to be start sites for RNA synthesis^{4,23–26}, α -amylase production would appear to be controlled by selecting specific transcriptional start sites. Different overlapping transcription units in adenovirus (1b and protein IX, and region 3 and P16) are also regulated by choosing the proper start sites (ref. 41; Fig. 3), as are the induced and constitutive forms of yeast invertase¹⁵¹.

Transcriptional control dependent on DNA rearrangement. Such interest has attended the possibility that rearrangements affect gene control that the proved examples of such rearrangements must at least be mentioned. Immunoglobulin formation, the variation in surface glycoproteins of trypanosomes, switching of the yeast mating-type locus and promoter insertion brought about by retrovirus movement in chromosomes all appear to be cases where gene activation, probably transcriptional activation, depends on DNA rearrangement. Actual transcriptional rate studies have not been done in most of these cases (see ref. 84).

RNA processing and its control

Although transcriptional control may be the predominant level of gene control, it is not the only one. To help understand post-transcriptional control, it is best first to consider more fully the steps in mRNA manufacture (see Fig. 1).

The sequence of events that convert a primary RNA transcript to a mRNA are collectively known as 'RNA processing'. These events begin very soon after initiation of an RNA chain. Before the new RNA chain is longer than 50 nucleotides^{34,85}, and perhaps before the RNA polymerase II moves past the initiation site, a 'cap', a m^7Gp residue⁷, is added to the first nucleotide (N). The middle phosphate of the m^7GpppN cap structure is not derived from the GTP but from the pppN that starts the RNA chain^{86–88}, strong evidence of RNA chain initiation at the cap site. Few, if any, polymerase II products escape this 5' end addition; even prematurely terminated RNA chains^{33,34,85} and chains synthesized *in vitro* by cell extracts supplemented with RNA polymerase II²⁵ are capped.

The poly(A) segment is added post-transcriptionally to heterogeneous nuclear RNA (hnRNA) in the cell nucleus^{8,9,20}. Although the precision of termination sites is not yet understood for polymerase II transcription, the enzyme has been shown to transcribe past the site for poly(A) addition in at least three adenovirus transcription units^{1,10–12}, both early and late SV40 (refs 13, 94) and polyoma transcription units^{95,96}, the β -globin transcription unit in mouse cells⁴ and chicken cells^{78,79}, and probably the transcription unit for the heavy chains of immunoglobulins^{97–99}. No clear cases of termination at the poly(A) site

have been documented by studying briefly labelled nuclear RNA. Thus addition of poly(A) may always require endonucleolytic cleavage and terminal addition of poly(A) to the cleaved 3' end of a nuclear RNA molecule. At least part of the signal for poly(A) addition is the AAUAAA sequence found 10–25 bases upstream from the poly(A) addition site^{100,101}.

Poly(A) addition usually, if not always, precedes in time the final, splicing^{15–17} steps of mRNA processing^{10,102–104}. Splicing, however, can occur in the absence of poly(A) addition¹⁰⁵ and mRNA formed during inhibition of poly(A) synthesis can at least in part enter polyribosomes¹⁰⁶, although it does not accumulate there as stable mRNA. These recent results suggest that the major (or perhaps sole) role of the poly(A) is cytoplasmic stabilization of the mRNA^{106–108} and that splicing for most transcripts might be an automatic event. Methylation of mRNA⁷, which is known to occur (on the average) to 1 out of 400 Ap residues in mammalian mRNAs^{109–111}, may occur before splicing, but its timing and role in mRNA formation remain speculative¹¹¹. Finally mRNAs leave the nucleus and enter the cytoplasm still bearing their poly(A) segment^{8,19}. With time, the poly(A) tail shortens⁸, which causes some mRNA sequences, >90% of which initially partition experimentally in the poly(A)-containing fraction, to fractionate in the poly(A)-lacking fraction^{92,93}. Extensive attempts with sea urchin material have failed to isolate any cloned DNA segment complementary to the mRNA molecules of the latter fraction that are not also present in the poly(A)-containing fraction (M. Nemer, personal communication) other than histone mRNA⁸⁹.

Transcription unit design and differential processing. Two categories of transcription units can be identified (Fig. 3). (1) Simple transcription units encode a single protein, whether or not processing of the primary transcript is required. Histone genes are simple transcription units that do not, in most cases, require either poly(A) addition or splicing⁹⁰. Protein IX mRNA of adenovirus¹¹² and the α -interferon mRNAs¹¹³ derive from simple transcriptional units whose products have poly(A) added but do not require splicing. The α - and β -globin transcription units are simple because, although their primary transcripts require poly(A) addition and splicing, only one functional mRNA is formed from the transcription unit^{114,115}.

(2) Complex transcription units are those whose primary transcript can give rise to two or more mRNAs that encode two different proteins. Two different structural arrangements are known for these units: two or more poly(A) sites and two or more variations in splicing patterns. Both virus and cell transcription units with multiple poly(A) sites are known—the major late adenovirus transcription unit^{10,63–65}, the heavy chains of immunoglobulins^{97–99} and the transcription unit encoding the peptide hormone calcitonin (R. Evans and G. Rosenfeld, personal communication). Thus far only virus transcription units are known that have one poly(A) site and two or more splicing sites—both transcription units of SV40 (refs 94, 117) and polyoma^{95,118}, most of the nine known adenovirus transcription units^{16,58,116,119} and many different retroviruses¹²⁰. A number of virus transcription units with multiple splicing arrangements can be integrated into cell DNA and still produce multiple mRNAs. Thus, it is clearly possible for DNA in the cell genome to encode primary transcripts that can be spliced in alternative ways. (Transcription units can have two or more poly(A) sites and variable splicing at one or both poly(A) sites, such as in the major late adenovirus transcription unit, Fig. 2.)

From these two types of complex transcription unit follow two general types of processing control⁵. (1) Differential processing: each transcript from a complex transcription unit must be processed in one of two or more ways. Several cases are known in which various poly(A) sites or splice sites can be chosen on the same primary transcript. (2) Process versus discard decisions: In this hypothetical possibility, there would be little or no processing of a particular transcript in one cell type or a cell in one particular physiological condition, but successful processing of the same primary transcript in another cell type

or another physiological condition (reviewed in ref. 5). Candidate transcripts of this type do exist in cultured cells and probably also in tissues but no specific case of gene control through process versus discard decisions is yet proved.

Adenovirus processing control. A differential choice of both poly(A) and splicing sites has been documented in the primary transcript of adenovirus mRNA, P16 (major late promoter; Fig. 2). Differential choice of poly(A) sites is also suspected to occur in at least one other transcription unit (region 2 transcripts that lead to mRNA formation with sequences in the 20–30 map unit region)^{58,121}.

Early in infection, in transcripts beginning at P16, the poly(A) sites at both 38.5 (L1) and 50 (L2) map units are used to generate poly(A)-containing nuclear RNA^{59–61}. About two L1 molecules are produced for each L2. Late in infection, in addition to the whole transcription unit being transcribed and the L1–5 being available for use, a change occurs so that L2 is used about twice as often as L1 (ref. 10). This change in ratio of poly(A) site usage accords with the change in amount of L1 versus L2 mRNAs in the cell nucleus. Late in infection, when five poly(A) sites are available, L3 is more frequently chosen than any other site but the ratio of L1–5 remains at 1:2:3:2:2 throughout the late course of infection (ref. 10 and J. R. Nevins, unpublished).

In addition to differential poly(A) choices, the early splicing pattern is different from the late one in the L1 poly(A)⁺ nuclear molecules. Examination of nuclear RNA indicates clearly that three products are made late in infection but only one major product early in infection⁶⁰. In addition, the initially labelled cytoplasmic molecules late in infection contain all three species of L1 mRNA whereas early in infection only a single species of labelled mRNA emerges in the cytoplasm after a brief label⁶⁰. Finally, the protein translated from the early L1 mRNA is encoded mainly, if not exclusively, in the 400 nucleotides that are retained in early L1 mRNA but are lost in the late mRNA⁶¹. Thus the evidence seems very strong that a change in splicing takes place in the handling of the same sequence of nucleotides contained in the nuclear poly(A)⁺ precursor molecule that is encoded in the 16–38.5 region on the adenovirus genome.

Immunoglobulin mRNA formation. Only two cases of differential processing are known for cellular mRNAs. During B-lymphocyte development and immunoglobulin synthesis, the μ heavy chain is first inserted as an integral protein in the plasma membrane^{97–99}. Later, a similar μ chain is found as part of secreted immunoglobulins. Analysis of cDNA clones of the two

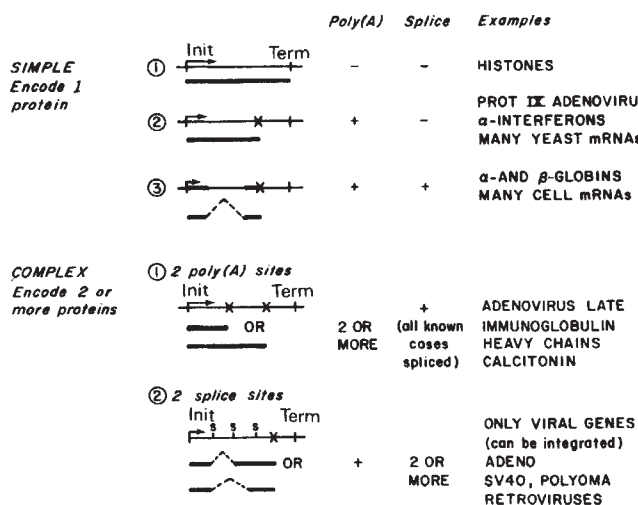


Fig. 3 Definition and examples of simple and complex transcription units. In each case, DNA is shown at the top and the mRNA(s) underneath. On the former are indicated sites of initiation (INIT.) and termination (TERM.) of transcription, splicing sites (s) and the position of the sequence which will become the site of poly(A) addition to the processed transcript.

mRNAs showed that the deduced carboxyl terminus of the μ membrane and μ secreted chains was different but the deduced amino termini were the same. Analysis of the genomic cloned DNA revealed two poly(A) sites in the same transcription unit and the primary transcript seems to cover both of them (ref. 97 and R. Wall, personal communication). Depending on the choice of poly(A) site two μ heavy chain mRNAs with different 5' most sequences can be formed, one for the μ membrane chain and one for the μ secreted chain. Further studies suggest that the same five exonic regions common to both μ chains also appear in δ chains⁹⁹. In this case, the 3' exonic component of the μ mRNA may be a consequence of choosing a poly(A) site more than 25 kb downstream. These studies were carried out on lymphocytic tumour cells in culture, but it is thought that a similar set of events occurs in the normal differentiation sequence in B cells.

Note that the molecular basis for none of these presumed differential choices of poly(A) and splice sites is known and it remains possible that template changes may dictate the post-transcriptional outcome. For example, rather than there being a single completed transcript in which a poly(A) site is chosen, a selective protein that bound to the DNA at one or the other poly(A) site might instead trigger the choice of poly(A) site during transcription.

Process versus discard decisions. Differential processing of nuclear RNA from either simple or complex transcription units might include the processing or discard of a particular nuclear RNA sequence^{5,18-20}. This type of processing is different from differential poly(A) choices or differential splicing choices where a transcript has a high probability of yielding one or another mRNA¹⁰. Although there is no evidence of variable process versus discard between two different cells, nuclear RNA turnover is substantial¹²² and nuclear RNA 'complexity' (average nucleotide content) is some 10-20 times greater than the complexity of the mRNA¹²³, perhaps greater than the difference expected on the basis of intron loss during nuclear RNA processing.

More recent analyses of nuclear RNA in cultured cells divide the nuclear RNA primary transcripts into two classes—those that produce mRNA sequences and those that do not. While all transcripts contain a cap³⁴ and >90% of newly capped mRNAs that enter polyribosomes (excluding histone mRNAs) contain poly(A)^{42,91}, poly(A) is added in the nucleus only to about one-quarter of the total primary transcripts⁹¹. Thus the rate of synthesis of caps is about four times that of poly(A). Moreover, when sequences complementary to eight cloned CHO cDNAs were examined in the poly(A)⁺ and poly(A)⁻ nuclear RNA fractions^{24,42}, whereas 70% of the sequences complementary to these clones were in the poly(A)⁺ fraction, only 20-25% of the capped molecules were in the poly(A)⁺ fraction. It therefore seems that capped poly(A)⁻ molecules whose sequences are not part of stable mRNAs make up a significant part of the nuclear RNA in cultured cells. Can these primary transcripts at some other time and place (embryonic cells? specifically differentiated cells?) be shifted from the poly(A)⁻ to the poly(A)⁺ category and then appear as stable mRNA^{5,20,125}? Specific clones complementary to these 'nucleus-confined' RNAs molecules are necessary to answer this possibility.

Cytoplasmic gene control

Too little is known about nuclear RNA transport to suggest the existence of differential transport. After the mRNA has reached the cytoplasm, however, the fates of mRNA molecules differ. Evidence exists for both a wide range of half lives for different specific mRNAs in the same cells^{62,92} and different half lives for the same mRNA in the same cell under different circumstances.

Of these possible control mechanisms the first would, of course, allow the cell to accumulate differential concentrations of two different mRNAs although the synthesis rates of the two

were identical, but its physiological significance remains obscure. Changes in the half life of specific mRNAs under different circumstances has attracted more attention. Two cases involving hormones are well documented.

Breast tissue can be explanted and continue to respond to hormonal stimulation by the production of casein mRNA and casein protein¹²⁶. When explants are cultured without prolactin for 48 h they lose 95% of their casein mRNA but can, when given prolactin, accumulate about 25,000 copies of the casein mRNA in 24 h. This accumulation is associated with an ~20-fold increase in the mRNA half life, but only a two- to threefold increase in the synthesis of casein mRNA sequences assayed by hybridization to casein cDNA.

A second reported instance of mRNA stabilization occurs with oestrogen treatment of roosters^{76,77}. As described earlier, mRNA sequences, for both VLDL and vitellogenin are synthesized at least 20-100-fold faster in liver nuclei from treated than from untreated animals. The rate of mRNA accumulation is maximal within about 6-8 h and from the curve of accumulation of vitellogenin and VLDL mRNAs during the first week of oestrogen treatment a half life of >24 h can be calculated. If, after the mRNA has reached a plateau value, the oestrogen is abruptly withdrawn, the vitellogenin and VLDL mRNA decline with a half life of ~3 h. Thus the half life changes from 24 to 3 h by removal of the hormone. Calculations of similar hormone-dependent changes in half life for steroid-induced hormones have also been reported¹²⁷, although the changing steroid levels were not so carefully controlled.

Differential half life of specific mRNAs also occurs during the growth cycle of adenovirus⁶². Transcriptional unit 1B produces two mRNAs—one 20S and one 14S. Early in infection the 20S form predominates while late in infection the 14S is predominant¹²⁹. The transcription rate of 1B sequences is constant from 5 to 24 h of infection⁶². When labelled nuclear RNA was examined early and late in infection no changes were noted, suggesting that differential splicing was not occurring⁶². However, cytoplasmic accumulations of 14S relative to 20S was greater late in infection than early in infection, pointing to an increased half life of the 14S mRNA. This mRNA appears to have the same cap site, the same splicing and poly(A) sites late in infection as it does early and thus the same mRNA has a different half life at two different times in the infectious cycle.

Finally, there are two established cases (and many others suspected) where preservation of a specific group of mRNAs occurs in the face of destruction of a large fraction of total cell mRNA¹³⁰⁻¹³². During erythroblast differentiation in mammals, the developing cell makes many different mRNAs including, in the final stages, globin mRNA. Globin RNA synthesis,

Table 1 Summary of frequency of various types of control

	Examples proved or strongly suggested	Possible
Nuclear		
Transcriptional		
Initiation	Many, >100	
Termination		
Premature (attenuation) readthrough	1	+
Nuclear RNA processing		
Poly(A) choice	~3	
Splicing choice	1	
Process versus discard		+
Cytoplasmic		
mRNA stability	~5 Specific, many general	
mRNA translation efficiency	~10 Specific, many general	

For those cases of control where only one or a few cases are proved, it is anticipated that these numbers will increase. It is not possible at present to even guess at the true frequency with which various mechanisms will be found ultimately to be used.

however, represents a tiny fraction of total nuclear RNA synthesis. During the final four or so cell divisions the cell becomes a highly specialized reticulocyte synthesizing over 90% globin plus a limited number of other specific proteins. These differentiative steps appear to represent conservation of the globin (and a few other erythrocytic) mRNAs and specific destruction of all other mRNAs¹³⁰⁻¹³². A second case of specific mRNA destruction has been described during development of the slime mould *Dictyostelium discoideum*. Cells which have aggregated will, if left alone, form a stalk, fruiting bodies and spores. This is accomplished with the aid of a set of new RNAs formed after aggregation. If the developmental cycle is interrupted after aggregation, the new developmentally regulated mRNAs are specifically destroyed¹³³.

Control of translation

Differential translation of finished mRNA molecules that are apparently already in the cell cytoplasm is recognized early in embryonic development, both for the average rate of protein synthesis per ribosome and for specific protein species. Unfertilized sea urchin eggs become at least 50-fold more active in protein synthesis from the same number of mRNAs after fertilization than before⁴. By contrast, in clam eggs the total rate of protein synthesis is unchanged after fertilization; however two-dimensional gel analysis of proteins newly formed *in vitro* compared with translation *in vitro* reveals certain proteins are formed specifically only pre- or post-fertilization¹³⁵ although the mRNAs are present.

Major shifts in total protein synthesis have also been recognized in non-embryonic cells. Cultured mammalian cells^{136,137} and insect cells¹³⁸ decrease total protein synthesis after a heat shock while the translation of certain 'heat-shock' proteins is probably enhanced. Cultured mammalian cells undergo a three-fold decrease in protein synthesis during each mitosis¹³⁹; and in virus-infected cells the synthesis of host proteins can be interrupted in several different ways although the mRNA is still present¹⁴⁰⁻¹⁴². It is also possible¹⁴³, though unproved, that tissue-specific favouritism for the translation of specific mRNAs or classes of mRNA occurs.

Control after translation

Several types of post-translational control can be envisioned that would affect the final concentration of particular proteins in given cells. For example, as polypeptides exist that can be cut by proteases into different, overlapping polypeptides with different activities, control of proteolysis represents a possible level of gene control. This apparently occurs in the posterior pituitary where cells can produce either adrenocorticotropin or β -lipotropin presumably from the same primary translation product^{144,145}.

Different fates for the same polypeptide, such as differential protein turnover or differential prosthetic group additions,

could obviously alter the enzymatic profile of a particular cell. Whether such differential post-translational events occur or are common remains to be explored.

Conclusion

The evidence summarized in this article demonstrates that a variety of mechanisms determine the abundance of specific mRNAs in eukaryotic cells, with the major level of control usually being transcription in the cell nucleus. Although the mechanism of the specific increases or decreases in transcription is not the subject of this review, it is widely believed that site-specific binding proteins either change chromatin structure or directly facilitate or prevent RNA polymerase attachment at specific sites on the DNA²¹. Only a few cases of supposed transcriptionally active proteins are known in vertebrate cells or virus-infected cells but quite a number of regulatory loci in lower eukaryotes, for example yeast, are now identified genetically (see, for example, refs 146-148). As yeasts have mRNAs with short half lives¹⁴⁹, most regulation of their rates of protein synthesis is probably due to controlled rates of transcription. Progress in the isolation of specific, transcriptionally active proteins in yeasts can be expected soon.

In addition to transcriptional initiation, a change in termination has apparently been identified in one adenovirus transcription unit. It is unknown whether differential termination will be a common mechanism for regulation of transcription, although evidence of frequent premature terminations^{33,34} is tantalizing in this regard. Note that transcriptional control for a particular gene does not preclude other levels of control.

The second level of nuclear gene control is processing. It is now clear that differential poly(A) choices and differential splicing can occur, but if a transcription unit has only one poly(A) site and only one possible spliced mRNA, then differential processing of the primary transcripts cannot occur. Of great importance is the possibility of process versus discard control. (Cloning of poly(A)⁻ primary transcripts from one cell source will be necessary to determine whether such transcripts ever become poly(A)⁺ transcripts in other cells.)

Differential mRNA turnover seems fairly common as even at this early stage in the study of specific cell mRNAs, several clear-cut cases of differential turnover are known.

Although multiple levels of gene control seem to exist, we are almost completely ignorant of the logic of gene control systems in eukaryotic cells or even whether most genes are ever subject to specific control. Regulation at transcription would appear for many genes to be quite strict; regulation of mRNA turnover, in contrast, is probably much less strict. What might distinguish genes that are tightly regulated from those that are not remains unknown. Finally the mechanisms and molecules that execute the various types of control remain unknown. Perhaps only by examining these regulatory molecules will the molecular logic of eukaryotic gene regulation become clear.

- Weber, J., Jelinek, W. & Darnell, J. E. *Cell* **10**, 611-616 (1977).
- Goldberg, S., Weber, J. & Darnell, J. E. *Cell* **10**, 617-621 (1977).
- Goldberg, S., Schwartz, H. & Darnell, J. E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4520-4523 (1977).
- Evans, R. M. *et al. Cell* **12**, 733-739 (1977).
- Hofer, E. & Darnell, J. E. *Cell* **23**, 585-593 (1981).
- Darnell, J. E. *Prog. Nucleic Acid Res. molec. Biol.* **22**, 327-353 (1979).
- Ziff, E. B. *Nature* **287**, 491-499 (1980).
- Shatkin, A. J. *Cell* **9**, 645-653 (1976).
- Brawerman, G. *CRC Rev. Biochem.*, 1-38 (1981).
- Jelinek, W. *et al. J. molec. Biol.* **75**, 515-532 (1973).
- Nevins, J. R. & Darnell, J. E. *Cell* **15**, 1477-1493 (1978).
- Fraser, N. W., Nevins, J. R., Ziff, E. & Darnell, J. E. *J. molec. Biol.* **129**, 643-656 (1979).
- Nevins, J. R., Blanchard, J. M. & Darnell, J. F. *J. molec. Biol.* **144**, 377-386 (1980).
- Ford, J. P. & Hsu, M. T. *J. Virol.* **28**, 795-801 (1978).
- Gilbert, W. *Nature* **271**, 501 (1978).
- Berget, S. M., Moore, C. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3171-3175 (1977).
- Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. *Cell* **12**, 1-8 (1977).
- Klessig, D. F. *Cell* **12**, 9-23 (1977).
- Scherrer, K. & Marcaud, L. *J. cell. Physiol.* **72**, Suppl. 1, 181-212 (1968).
- Darnell, J. E., Jelinek, W. & Molloy, G. *Science* **181**, 1215-1222 (1973).
- Korn, L. J. & Gurdon, J. B. *Nature* **289**, 461-465 (1981).
- Weisbrod, S. *Nature* **297**, 289-295 (1982).
- Roeder, R. in *RNA Polymerases* (eds Losick, R. & Chamberlin, M.) 285-329 (Cold Spring Harbor Laboratory, New York, 1976).
- Corden, J. *et al. Science* **209**, 1406-1414 (1980).
- Ziff, E. & Evans, R. M. *Cell* **15**, 1463-1475 (1978).
- Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. *Cell* **18**, 469-484 (1979).
- Baker, C. & Ziff, E. *J. molec. Biol.* **149**, 189-221 (1981).
- Goldberg, M. thesis, Stanford Univ. (1978).
- Benoist, C. & Chambon, P. *Nature* **290**, 304-310 (1981).
- Ucker, D. S., Ross, S. R. & Yamamoto, K. R. *Cell* **27**, 257-266 (1981).
- Struhl, L. & Davis, R. *J. molec. Biol.* **52**, 553-568 (1970).
- Struhl, K. *J. molec. Biol.* **52**, 569-576 (1970).
- McKnight, S. L., Gavis, E. R., Kingsbury, R. & Axel, R. *Cell* **25**, 385-398 (1981).
- Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. *Cell* **27**, 279-288 (1981).
- Fraser, N. W., Sehgal, P. B. & Darnell, J. F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2571-2575 (1979).
- Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S. & Darnell, J. E. *Cell* **19**, 69-78 (1980).
- Jacob, F. & Monod, J. *J. molec. Biol.* **3**, 318-356 (1961).
- Miller, J. H. & Reznikoff, W. W. (eds) *The Operon* (Cold Spring Harbor Laboratory, New York, 1978).
- Yanofsky, C. *Nature* **289**, 751-759 (1981).
- Hayashi, M., Spiegelman, S., Franklin, N. C. & Lauria, S. E. *Proc. natn. Acad. Sci. U.S.A.* **49**, 729-736 (1963).
- Bertrand, K., Squires, C. & Yanofsky, C. *J. molec. Biol.* **103**, 319-337 (1976).
- Lozeron, H. A., Dahlberg, J. E. & Szybalski, W. *Virology* **71**, 262-277 (1976).
- Harpold, M. M., Evans, R. M., Salditt-Georgieff, M. & Darnell, J. E. *Cell* **17**, 1025-1035 (1979).
- Salditt-Georgieff, M. & Darnell, J. E. *Molec. cell. Biol.* **3** (in the press).
- Cox, R. F. *Cell* **7**, 455-465 (1976).

45. Wall, R., Philipson, L. & Darnell, J. E. *Virology* **50**, 27-34 (1972).
46. Weber, J. thesis, Rockefeller Univ. (1979).
47. Derman, E. *et al. Cell* **23**, 731-739 (1981).
48. Lowenhaupt, K., Trent, C. & Lingrel, J. B. *Dev. Biol.* **63**, 441-454 (1978).
49. Nevins, J. R., Ginsberg, H. W., Blanchard, J. M., Wilson, M. C. & Darnell, J. E. *J. Virol.* **32**, 727-733 (1979).
50. Carter, T. H. & Blanton, R. J. *Virol.* **25**, 664-674 (1978).
51. Pettersson, U. & Matthews, M. B. *Cell* **12**, 741-750 (1977).
52. Wilson, M. C., Fraser, N. W. & Darnell, J. E. *Virology* **94**, 178-184 (1979).
53. Fraser, N. W. & Ziff, E. *J. molec. Biol.* **155** (in the press).
54. Jones, N. & Shenk, T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3665-3669 (1979).
55. Harrison, T., Graham, F. & Williams, J. *Virology* **77**, 319-329 (1977).
56. Berk, A. J., Lee, F., Harrison, T., Williams, J. & Sharp, P. A. *Cell* **17**, 935-944 (1979).
57. Nevins, J. R. *Cell* **26**, 213-220 (1981).
58. Shenk, T., Jones, N., Colby, W. & Fowlkes, D. *Cold Spring Harb. Symp. quant. Biol.* **44**, 367-375 (1979).
59. Chow, L. T., Broker, T. R. & Lewis, J. B. *J. molec. Biol.* **134**, 265-303 (1979).
60. Shaw, A. R. & Ziff, E. *Cell* **22**, 905-916 (1980).
61. Nevins, J. R. & Wilson, M. C. *Nature* **290**, 113-118 (1981).
62. Akusjarvi, G. & Persson, H. *Nature* **292**, 420-426 (1981).
63. Wilson, M. C. & Darnell, J. E. *J. molec. Biol.* **148**, 231-251 (1981).
64. Ziff, E. & Fraser, N. W. *J. Virol.* **25**, 897-906 (1978).
65. McGrogan, M. & Raskas, H. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 625-629 (1978).
66. Nevins, J. R. & Darnell, J. E. *J. Virol.* **25**, 811-823 (1978).
67. Nevins, J. R. & Jensen-Winkler, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1893-1897 (1980).
68. Acheson, N. H. in *Molecular Biology of Tumor Viruses* 2nd edn, Pt 2 (ed. Tooze, J.) (Cold Spring Harbor Laboratory, New York, 1980).
69. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
70. Rio, D., Robbins, A., Meyers, R. & Tjian, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5706-5710 (1980).
71. Myers, R. M., Rio, D. C., Robbins, A. K. & Tjian, R. *Cell* **25**, 373-384 (1981).
72. Gutai, M. W. & Nathans, D. *J. molec. Biol.* **126**, 259-274 (1978).
73. Swaneck, G. E., Nordstrom, J. L., Krenzlee, F., Tsai, M.-J. & O'Malley, B. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1049-1053 (1979).
74. McKnight, G. S. & Palmiter, R. D. *J. biol. Chem.* **254**, 9050-9058 (1978).
75. Tsai, S. *et al. Biochemistry* **17**, 5733-5780 (1979).
76. Goldberger, R. in *The Role of RNA in Development and Reproduction* (eds Niu, M. & Chuang, H. H.) 215-231 (Science, Beijing, 1981).
77. Wiskocil, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4474-4478 (1980).
78. Groudine, M., Peretz, M. & Weintraub, H. *Molec. cell. Biol.* **1**, 281-288 (1981).
79. Groudine, M. & Weintraub, H. *Cell* **24**, 333-334 (1981).
80. Finlayson, J. S., Asofsky, R., Potter, M. & Runner, C. *Science* **149**, 981-982 (1965).
81. Derman, E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5425-5429 (1981).
82. Hagenbuchle, O. *et al. Nature* **289**, 643-646 (1981).
83. Young, R. A., Hagenbuchle, O. & Schibler, U. *Cell* **23**, 451-458 (1981).
84. *Cold Spring Harb. Symp. quant. Biol.* **46** (1981).
85. Babich, A., Nevins, J. R. & Darnell, J. E. *Nature* **287**, 246-248 (1980).
86. Contreras, R. & Fiers, W. *Nucleic Acids Res.* **9**, 215-236 (1981).
87. Venkatesan, S. & Moss, B. *J. J. biol. Chem.* **255**, 2835-2842 (1980).
88. Venkatesan, S., Gershowitz, H. & Moss, B. *J. J. biol. Chem.* **255**, 2829-2834 (1980).
89. Adesnik, M. & Darnell, J. E. *J. molec. Biol.* **67**, 397-406 (1972).
90. Hentschel, C. C. & Birnstiel, M. L. *Cell* **25**, 30-313 (1981).
91. Salditt-Georgieff, M., Harpold, M., Wilson, M. & Darnell, J. E. *Molec. cell. Biol.* **1**, 179-187 (1981).
92. Harpold, M., Wilson, M. & Darnell, J. E. *Molec. cell. Biol.* **1**, 188-198 (1981).
93. Minty, A. J. & Gros, F. *J. molec. Biol.* **139**, 61-83 (1980).
94. Reddy, V. B. *et al. Science* **200**, 494-502 (1978).
95. Legon, S., Flavell, A., Cowie, A. & Kamen, R. *Cell* **16**, 373-388 (1979).
96. Acheson, N. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4754-4758 (1978).
97. Rogers, J. *et al. Cell* **20**, 303-312 (1980).
98. Alt, F. W. *et al. Cell* **20**, 293-301 (1980).
99. Maki, R. *et al. Cell* **24**, 353-365 (1981).
100. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211-214 (1976).
101. Fitzgerald, M. & Shenk, T. *Cell* **24**, 251-260 (1981).
102. Lai, C.-J., Dhar, R. & Khoury, G. *Cell* **14**, 971-982 (1978).
103. Tilghman, S. M., Curtis, P. J., Tiemeier, D. C., Leder, P. & Weissman, C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1309-1313 (1978).
104. Roop, D. R. *et al. Cell* **15**, 671-685 (1978).
105. Zeevi, M., Nevins, J. R. & Darnell, J. E. *Cell* **26**, 39-46 (1981).
106. Zeevi, M., Nevins, J. R. & Darnell, J. E. *Molec. cell. Biol.* **3** (in the press).
107. Nudel, U. *et al. Eur. J. Biochem.* **64**, 115-125 (1976).
108. Huez, G., Bruck, C. & Cleuter, Y. *Proc. natn. Acad. Sci. U.S.A.* **78**, 908-911 (1981).
109. Salditt-Georgieff, M. *et al. Cell* **7**, 227-237 (1976).
110. Sommer, S. *et al. Nucleic Acids Res.* **3**, 749-766 (1976).
111. Chen-Kiang, S., Nevins, J. R. & Darnell, J. E. *J. molec. Biol.* **135**, 733-752 (1979).
112. Alestrom, P. G. *et al. Cell* **19**, 671-681 (1980).
113. Nagata, S., Mantei, N. & Weissman, C. *Nature* **287**, 401-408 (1980).
114. Konkel, D. A., Tilghman, S. M. & Leder, P. *Cell* **15**, 1125-1132 (1978).
115. Nishioka, Y. & Leder, P. *Cell* **18**, 875-882 (1979).
116. Kitchingman, G., Lai, S. & Westphal, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4392-4395 (1977).
117. Fiers, W. *et al. Nature* **273**, 113-120 (1978).
118. Soeda, E., Arrand, J. R., Smolar, N., Walsh, J. E. & Griffin, B. E. *Nature* **283**, 445-453 (1980).
119. Berk, A. J. & Sharp, P. A. *Cell* **14**, 695-711 (1978).
120. *Cold Spring Harb. Symp. quant. Biol.* **44**, Pt 2 (1979).
121. Stillman, B. W., Lewis, J. B., Chow, L. T., Matthews, M. B. & Smart, J. E. *Cell* **23**, 497-508 (1981).
122. Soeiro, R., Vaughn, M. H., Warner, J. R. & Darnell, J. E. *J. Cell Biol.* **39**, 112-118 (1968).
123. Lewin, B. (ed.) in *Gene Expression* 2nd edn, 748 (Wiley-Interscience, New York, 1980).
124. Ventateson, S., Nakazato, H., Kopp, D. W. & Edmonds, M. *Nucleic Acids Res.* **6**, 1097-1110 (1979).
125. Davidson, E. H. & Britten, R. J. *Science* **204**, 1052-1059 (1979).
126. Guyette, W. A., Matuski, R. J. & Rosen, J. M. *Cell* **17**, 1013-1023 (1979).
127. Greenberg, J. R. *Nature* **240**, 102-104 (1972).
128. Palmiter, R. J. *biol. Chem.* **248**, 8260 (1973).
129. Spector, D. J., McGrogan, M. & Raskas, H. J. *J. molec. Biol.* **126**, 395-414 (1978).
130. Aviv, H., Volloch, Z., Bastos, R. & Levy, S. *Cell* **84**, 495-503 (1976).
131. Bastos, R. N., Volloch, Z. & Aviv, A. *J. molec. Biol.* **110**, 191-203 (1977).
132. Volloch, V. & Hausman, D. in *Organization and Expression of Globin Genes* (eds Stamatoyannopoulos, G. & Nienhaus, A. W.) 251-268 (Liss, New York, 1981).
133. Chung, S., Landfear, S. M., Blumberg, D. D., Cohen, N. S. & Lodish, H. F. *Cell* **24**, 785-797 (1981).
134. Brandhoest, B. P. *Dev. Biol.* **52**, 310-317 (1976).
135. Rosenthal, E., Hunt, T. & Roderman, J. V. *Cell* **20**, 487-494 (1980).
136. McCormick, W. & Penman, S. *J. molec. Biol.* **39**, 315-333 (1969).
137. Kelley, P. M. & Schlessinger, M. J. *Cell* **15**, 1277-1286 (1978).
138. Storti, R. V., Scott, M. P., Rich, A. & Pardue, M. L. *Cell* **22**, 825-834 (1980).
139. Fan, H. & Penman, S. *J. J. molec. Biol.* **50**, 665-670 (1970).
140. Trachsel, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 770-774 (1980).
141. Skup, D., Zarbl, H. & Millward, S. *J. molec. Biol.* **151**, 35-56 (1981).
142. Fernandez-Munoz, R. & Darnell, J. E. *J. Virol.* **18**, 719-726 (1976).
143. Heywood, S. M., Kennedy, D. S. & Bester, A. J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2428-2431 (1974).
144. Roberts, J. L. & Herbert, E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4826-4830, 5300-5304 (1977).
145. Mains, R. E. & Epper, B. A. *J. biol. Chem.* **251**, 4115-4120 (1976).
146. Bach, M.-L., Lacroute, F. & Botstein, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 386-390 (1979).
147. Klar, A. J. S., Strathern, J. N., Broach, J. R. & Hicks, J. B. *Nature* **289**, 239-244 (1981).
148. Broach, J. R. *J. molec. Biol.* **131**, 41-53 (1979).
149. Hutchison, H. T., Hartwell, L. H. & McLaughlin, C. S. *J. Bact.* **99**, 807-814 (1969).
150. Hofer, E., Hofer-Warbinek, R. & Darnell, J. E. *Jr Cell* (in the press).
151. Carlson, M. & Botstein, D. *Cell* **28**, 145-154 (1982).

ARTICLES

Implications for trace gases and aerosols of large negative ion clusters in the stratosphere

F. Arnold, A. A. Viggiano & H. Schlager

Max-Planck-Institut für Kernphysik, Heidelberg, FRG

Stratospheric ions are of considerable interest because they not only control atmospheric electrical properties but also may influence stratospheric trace gases and aerosols. The first detailed in situ composition measurements of the stratospheric heavy negative ion clusters are now reported; these measurements have interesting implications for trace gases and aerosols.

IN SITU composition measurements of stratospheric ions became technically possible only during recent years. As found from the first *in situ* negative ion composition measurements^{1,2}, stratospheric sulphuric acid vapour has a strong influence on stratospheric negative ions. Thus, because H₂SO₄ vapour is the most important condensable vapour in the stratosphere, and influences the formation of the stratospheric aerosol layer³,

negative ion composition measurements can also yield information on aerosols. In particular HSO₄⁻(H₂SO₄)_n(HNO₃)_m ions have been found⁴ to be abundant at altitudes above ~27 km, whereas NO₃⁻(HNO₃)_n ions are dominant below this height. It has been proposed¹ that the former ions are formed from the latter ones by reactions involving H₂SO₄ vapour followed by H₂SO₄-association. This suggestion has received strong support

from studies of relevant ion reactions conducted by Viggiano *et al.*⁵. Attention has also been drawn⁶ to the possibility of inferring stratospheric H₂SO₄-vapour concentrations from *in situ* negative ion composition measurements.

Recent *in situ* negative ion composition measurements conducted by the Heidelberg⁷⁻⁹ and Bruxelles¹⁰ groups confirmed the original suggestions, but also provided strong evidence for the presence of very massive negative ion clusters in the region above 27 km. In fact, it appears that most negative ions fall outside the mass range (≤ 320 AMU), accessible to previous *in situ* ion composition measurements in which mass resolving modes of operation were used. The so called 'total throughput modes' allowed detection of heavy negative ion clusters without yielding detailed mass information. We have investigated large negative ion clusters in the stratosphere and their bearing on aerosol formation^{11,12}. Our detailed *in situ* measurements are given below.

Negative ion composition measurements

The principle of the balloon-borne ion mass spectrometer used in the present studies has been discussed in detail elsewhere^{1,4,12} and therefore will not be reviewed here. Note, however, that the present instrument has an extended mass range covering 0–612 AMU, which was achieved by reducing the frequency of the R.F.-generator of the quadrupole mass filter from 2.0 to 1.4 MHz.

The improved instrument was flown on a balloon during daytime on 17 October 1981 over southwestern France (geographical coordinates of the measuring site: 44° N, 0° W) and reached a maximum altitude of 34 km. The composition measurements of heavy negative ions were made between 32 and 34 km during the ascent and float portions of the balloon flight between 09.24 and 09.47 LT.

Typical negative ion mass spectra obtained in this height region are shown in Fig. 1. The mass resolution used was only

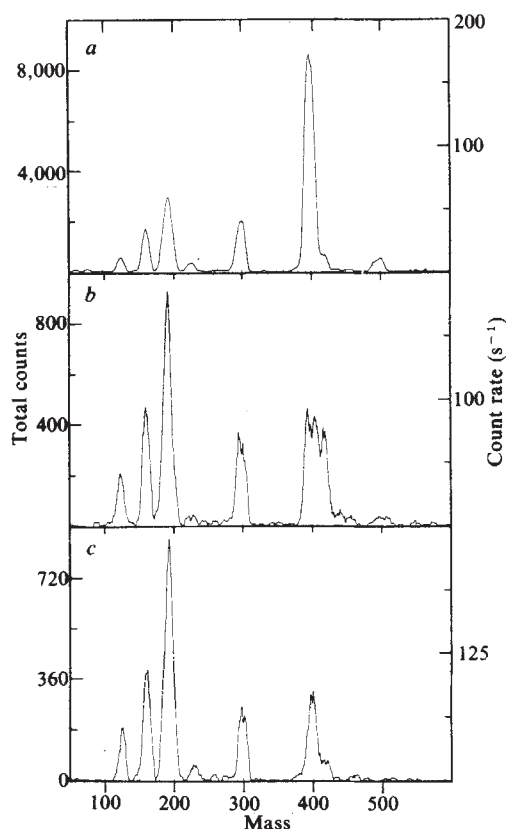


Fig. 1 Negative ion mass spectra measured at altitudes of 34 (a), 32.9 (b) and 32.4 (c) km. Total integration times were 1,000, 140 and 100 s. A running window of 51 channels was used to improve statistics.

moderate (average peak width at half maximum equal to about 14 AMU) to achieve a high sensitivity for ion detection, particularly at large masses. Due to the large peak width, masking of small peaks may have occurred. The estimated uncertainty for mass identification is about ± 1 AMU large peaks and ± 2 or 3 AMU for smaller peaks depending on the ion pulse statistics, which, in turn, depend on the integration time. Due to a much larger integration time, the peaks of the spectrum obtained at 34 km (float altitude) have much better defined shapes.

Results for the spectral region 0–320 AMU are rather similar to those obtained previously^{1,2,4,7,10}. Therefore, we will focus on the mass region 320–612 amu, which was not previously accessible.

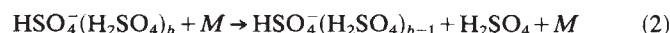
Clearly, this region contains major peaks, which is particularly striking in Fig. 1a, where the dominant mass peak is 391 ± 1 . A compilation of the observed peaks is given in Table 1 together with fractional count rates and tentative ion identifications. No attempt has been made to convert ion count rates to ion abundances as this conversion is complicated by cluster ion break-up that occurs during ion sampling. As has been discussed previously⁴, break-up should not drastically alter the ion composition at these heights. However, cluster ions may have lost one weakly bonded ligand molecule during sampling.

Negative ion chemistry

Table 1 reveals that the major heavy ions have HSO₄⁻ cores containing mainly H₂SO₄ ligands. Size distributions for these ions are shown in Fig. 2. The distributions are relatively flat up to $b = 3$ (b is the number of ligands) and fall off steeply for larger b . Most likely, these ions are formed by



followed by further displacement of HNO₃ by H₂SO₄ and finally H₂SO₄ association. As b increases the thermal dissociation process



becomes important and leads to the observed fall-off of the size distribution.

The predicted size distribution can be obtained from a steady-state treatment yielding

$$\frac{[\text{HSO}_4^-(\text{H}_2\text{SO}_4)_{b+1}]}{[\text{HSO}_4^-(\text{H}_2\text{SO}_4)_b]} = \left(1 + \frac{\alpha n_+}{k_+[M][\text{H}_2\text{SO}_4]} + \frac{k[M]}{k_+[M][\text{H}_2\text{SO}_4]} \right)^{-1} \quad (3)$$

where k_+ , k and α are the rate coefficients for ternary H₂SO₄-association, thermal dissociation and ion-ion recombination, respectively. The quantities $[M]$ and n_+ are the total gas and positive ion concentrations. Taking the phenomenological binary rate coefficient, $k_2 = k_+[M]$, for H₂SO₄-association equal to $10^{-9} \text{ cm}^3 \text{ s}^{-1}$ (saturated ternary process) and a sulphuric acid vapour concentration $[\text{H}_2\text{SO}_4] = 3 \times 10^6 \text{ cm}^{-3}$ (derived from the abundance ratio of measured HSO₄⁻ and NO₃⁻-ion clusters using the method described in ref. 6) one finds that the second term on the right-hand side of equation (3) becomes $\ll 1$. Consequently, if H₂SO₄-association proceeds close to the collision rate, as expected, the ion abundance ratio (3) becomes ≈ 1 when thermal dissociation is negligible. As seen from Fig. 2, this seems approximately to be the case for b up to 2–3. When thermal dissociation becomes very efficient, the third term on the right-hand side of equation (3) becomes dominant and the predicted ion abundance ratio becomes $\ll 1$. This seems to be the case for $b \geq 3$.

From Table 1, one sees that there are less abundant heavy ions containing HNO₃, H₂O and probably HSO₃ ligands. While H₂O and HNO₃ ligands were expected, the presence of HSO₃ ligands comes as a surprise. H₂O, HNO₃ and HSO₃ probably associate to the negative ion clusters and are lost by thermal dissociation, displacement by more strongly bonded ligands or

Table 1 Mass numbers of major negative ions as measured at 32.5, 32.9 and 34.0 km altitude

Mass	Fraction	.006	Ion
62 ± 2	0.006	NO ₃ ⁻	
125 ± 1	0.065	NO ₃ ⁻ · HNO ₃	
143 ± 2	0.009	NO ₃ ⁻ · HNO ₃ · H ₂ O	
160 ± 1	0.140	HSO ₄ ⁻ · HNO ₃	
178 ± 2	0.020	HSO ₄ ⁻ · HNO ₃ · H ₂ O	
188 ± 1	0.190	NO ₃ ⁻ (HNO ₃) ₂	
195 ± 3	0.100	HSO ₄ ⁻ · H ₂ SO ₄	
223 ± 1	0.018	HSO ₄ ⁻ (HNO ₃) ₂	
251 ± 2	0.003	NO ₃ ⁻ (HNO ₃) ₃	
276 ± 3	0.010	HSO ₄ ⁻ H ₂ SO ₄ · HSO ₃	
293 ± 1	0.120	HSO ₄ ⁻ (H ₂ SO ₄) ₂	
311 ± 3	0.003	HSO ₄ ⁻ (H ₂ SO ₄) ₂ H ₂ O	
374 ± 3	0.003	HSO ₄ ⁻ (H ₂ SO ₄) ₂ HSO ₃	
391 ± 2	0.140	HSO ₄ ⁻ (H ₂ SO ₄) ₃	
409 ± 3	0.130	HSO ₄ ⁻ (H ₂ SO ₄) ₃ H ₂ O	
427 ± 3	0.022	HSO ₄ ⁻ (H ₂ SO ₄) ₃ (H ₂ O) ₂	
454 ± 2	0.015	HSO ₄ ⁻ (H ₂ SO ₄) ₃ HNO ₃	
472 ± 3	0.004	HSO ₄ ⁻ (H ₂ SO ₄) ₃ HSO ₃	
489 ± 3	0.012	HSO ₄ ⁻ (H ₂ SO ₄) ₄	
507 ± 3	0.015	HSO ₄ ⁻ (H ₂ SO ₄) ₄ H ₂ O	
587 ± 3	0.004	HSO ₄ ⁻ (H ₂ SO ₄) ₅	

Tentative ion identifications and fractional ion count rates for the 32.9 km spectrum are given.

possibly by reactions. The mass identity of the ligand we identify as HSO₃, however, does not exclude other possibilities. For example, the mass of an HNO₃ and H₂O molecule added together (81 AMU) is the same as that of HSO₃. The relative abundances of the ions HSO₄⁻(H₂SO₄)_nH₂O, HSO₄⁻(H₂SO₄)_nHNO₃ and HSO₄⁻(H₂SO₄)_n · 81, however, leads one to believe that mass 81 corresponds to only one ligand. HSO₃ is then the most likely candidate for this ligand because it is the only acid of this mass to be expected in the atmosphere and acids are the preferred negative ion ligands^{1,2,4}.

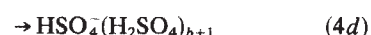
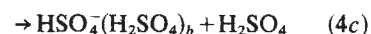
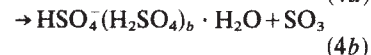
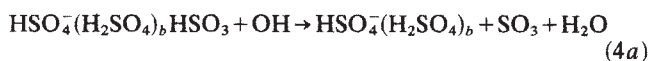
Because both, H₂O and HNO₃ are relatively abundant stratospheric trace gases (approximate number densities are 10¹² cm⁻³ and 10⁸ cm⁻³, respectively, at 30 km altitude), they associate rapidly to the negative ion clusters. However, both molecular species probably do not bond very strongly to the larger ion clusters and, therefore, may undergo rapid thermal detachment or displacement by more strongly bonded ligands, such as H₂SO₄. The ligand bond strength seems to increase in the sequence H₂O, HNO₃, H₂SO₄ for the cluster ion sizes considered here.

Measured abundance ratios for hydrate ions to their corresponding unhydrated forms are given in Table 2. Generally, the ratios decrease as the size of the unhydrated ion increases, in qualitative harmony with a simple charge-dipole interaction model. However, for HSO₄⁻(H₂SO₄)_b clusters, the ratios increase as *b* becomes >2. This behaviour, which as yet has not been observed for cluster ions, probably indicates strong cooperative bonding. In other words, ligand-ligand interaction seems to become important. In view of the large heat of mixing for bulk phase H₂SO₄-H₂O mixtures, this finding may qualitatively be expected. Quantitatively, however, it is surprising that cooperative effects become important already for relatively small ion cluster sizes.

The probable detection of HSO₃-ligands indicates that HSO₃-vapour is present in the stratosphere and associates to the negative ion clusters.

This view is supported by recent high pressure ion source studies made at our laboratory¹³. It was found that besides HSO₄⁻(H₂SO₄)_b, clusters of the type HSO₄⁻(H₂SO₄)_b · HSO₃ were most abundant in a gas mixture of SO₂ and room air (mostly N₂, O₂, and H₂O) ionized by a gas discharge. The HSO₃-ligand evidently bonds relatively strongly to HSO₄⁻(H₂SO₄)_b ion clusters. This is also supported by collision-induced fragmentation studies of HSO₄⁻(H₂SO₄)_bHSO₃ ions.

The fate of stratospheric HSO₃-ligands is, however, uncertain. They may detach thermally, become displaced by H₂SO₄, or undergo chemical reactions. The latter processes may involve reactions with OH-radicals

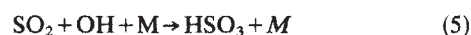


In process (4d) another ligand, for example H₂O, also contained in the precursor ion complex may become 'boiled off' carrying away the chemical energy released.

Processes (4a-d) would represent 'ion catalysed reactions'¹⁷.

Inferred HSO₃-abundance

Although the HSO₃-radical has not yet been detected in the stratosphere, its presence is expected. It is thought^{3,14,15} that HSO₃ arises from



and is lost preferentially by



whose rate coefficient is not known. The SO₃ radical is then presumed to react with H₂O to form H₂SO₄. Within the framework of their complex photochemical-diffusive model, Turco *et al.* have calculated the stratospheric HSO₃ abundance, yielding ~10⁵ cm⁻³ at 30 km altitude.

From the ion composition data reported above, a lower limit to the HSO₃ abundance can be inferred. Assuming that HSO₄⁻(H₂SO₄)_bHSO₃ ions are formed by HSO₃-association proceeding at the collision rate and that they are lost only by ion-ion recombination a steady-state treatment yields a lower limit of [HSO₃] = (2-10) × 10⁴ cm⁻³. When compared with this inferred value, the model calculation of Turco *et al.* are in approximate agreement.

As the model assumed reaction (6) to proceed at the collision rate (rate coefficient 10⁻¹¹ cm² s⁻¹) the above agreement may imply that reaction (6) is in fact fast. The HSO₃-lifetime compared with attachment to negative ions is ~10 times larger than the lifetime compared with reaction (6) if upper limits to the corresponding rate coefficients are used (ref. 16, and personal communications from D. Smith, N. G. Adams and E. Alge, and P. J. Crutzen). Consequently, the HSO₃ abundance would be of the order of 10⁶ cm⁻³ at 30 km if reaction (6) was unimportant and if the HSO₃-loss would proceed by ion attachment.

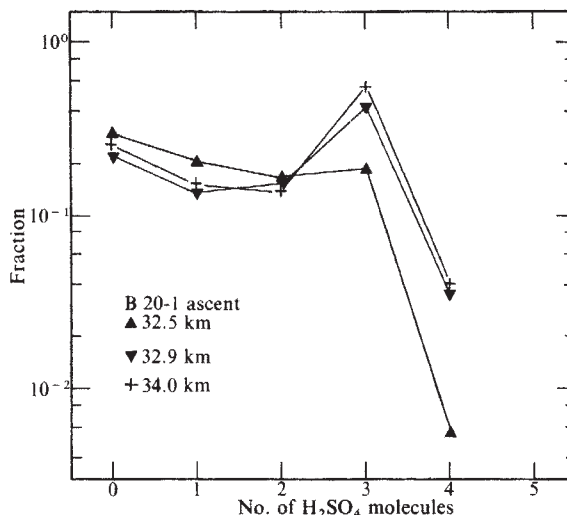


Fig. 2 Measured distributions of HSO₄⁻ clusters containing 0-5 H₂SO₄-ligands. Count rates were normalized to the total count rates for HSO₄⁻ clusters.

At lower altitudes, however, where the OH-abundance is lower and where the total ion concentration is larger, a significant fraction of HSO_3^- may be lost by ion reactions. Because, at these heights, $\text{NO}_3^-(\text{HNO}_3)_a$ ions are most abundant, another ion process involving HSO_3^- may be considered. It involves



followed by recombination of $\text{HSO}_4^-(\text{HNO}_3)_a$ with a positive ion, mostly $\text{H}^+(\text{CH}_3\text{CN})_l(\text{H}_2\text{O})_m$ and $\text{H}^+(\text{H}_2\text{O})_n$ possibly leading to the formation of an H_2SO_4 molecule. Reaction (7) is exothermic^{5,18,19} by $\sim 40\text{--}50 \text{ kcal mol}^{-1}$.

An important implication concerning analytical applications of *in situ* composition measurements of stratospheric negative ions is that the detection of HSO_4^- cores would not allow us unambiguously to infer H_2SO_4 abundances, thus the previously derived^{6,8,9} H_2SO_4 abundances should be reported as H_2SO_4 and HSO_3^- abundances below $\sim 27 \text{ km}$. Above this height, however, H_2SO_4 should be much more abundant than HSO_3^- and the derived quantity should refer only to H_2SO_4 . This conclusion is supported by the fact that above 27 km the negative ion clusters also contain H_2SO_4 ligands.

Thermodynamic properties of ion clusters

The present ion composition data can also be used to infer unknown thermodynamic properties of large negative ion clusters.

For this purpose, the concentrations of the clustering gas phase molecular species, H_2SO_4 , HNO_3 , HSO_3^- and H_2O as well as the gas temperature have to be known. Sulphuric acid concentrations were inferred independently from the negative ion composition (see above) whereas HNO_3^- , HSO_3^- and H_2O -concentrations were taken from model calculations. The actual values used are: $[\text{H}_2\text{SO}_4] = 3 \times 10^6 \text{ cm}^{-3}$; $[\text{HNO}_3] = 1 \times 10^8 \text{ cm}^{-3}$; $[\text{HSO}_3^-] = 1 \times 10^5 \text{ cm}^{-3}$; $[\text{H}_2\text{O}] = 1 \times 10^{12} \text{ cm}^{-3}$. The gas temperature was measured by various radiosonde stations (CNES Meteorological Service, personal communication) in the vicinity of the balloon launch site on the day of our experiment. In addition, the gas temperature can be inferred from the simultaneously measured positive ion composition with relatively high accuracy. The resulting temperature for $32\text{--}34 \text{ km}$ was consistently 233 K .

The above concentrations, with the exception of $[\text{HSO}_3^-]$, are sufficiently large to allow for clustering time scales less than the ion-ion recombination lifetime ($2,000 \text{ s}$). Consequently, quasi-equilibria for ion-clustering of these species should be established. For HSO_3^- clustering, this is not the case and therefore only lower limits to bond strengths can be inferred.

First, we will examine H_2SO_4 -clustering to $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_b$ ions ($b = 2, 3, 4$). Equilibrium constants, K , can easily be inferred from the measured ion abundance ratios and the derived H_2SO_4 abundance. The free energy difference, $\Delta G = -RT \ln K$, can then be evaluated using the measured gas temperature. The resulting ΔG values are listed in Table 3 along with free enthalpy differences, $\Delta H = \Delta G + T\Delta S$, estimated by assuming a typical common entropy difference²⁰, ΔS , of $25 \text{ cal mol}^{-1} \text{ K}^{-1}$.

The inferred ΔG -values for H_2SO_4 -clustering to $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_b$ clusters with $b = 3\text{--}4$ are both about 13 kcal mol^{-1} which is comparable to the heat of vaporization for bulk phase H_2SO_4 (ref. 21) ($12.6 \text{ kcal mol}^{-1}$ at 233 K). The inferred ΔG represents only a lower limit, because thermal dissociation of $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_3$ does not seem to be sufficiently fast to allow for an equilibrium between $b = 2$ and 3 to be established (see Fig. 2).

This result suggests that the properties of the observed heavy $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_b$ clusters ($b = 4$ and 5) already exhibit a trend towards bulk phase properties. Qualitatively, this may be expected as this trend is well known from laboratory studies^{20,22,23} of cluster ion equilibria such as $\text{H}^+(\text{H}_2\text{O})_n$, $\text{NO}_3^-(\text{H}_2\text{O})_n$ and $\text{Cl}^-(\text{SO}_2)_n$. It was found that ΔH -values for clustering

Table 2 Ratios of measured hydrate ion abundances to their unhydrated precursors

Unhydrated species	Ratio
$\text{NO}_3^-(\text{HNO}_3)_1$	0.14
$\text{HSO}_4^-(\text{HNO}_3)_1$	0.14
$\text{HSO}_4^-(\text{H}_2\text{SO}_4)_2$	0.03
$\text{HSO}_4^-(\text{H}_2\text{SO}_4)_3$	0.92
$\text{HSO}_4^-(\text{H}_2\text{SO}_4)_4$	1.25
$\text{HSO}_4^-(\text{H}_2\text{SO}_4)_3\text{H}_2\text{O}$	0.17

of a molecular species to ion approach the heat of vaporization of the molecular species as the number of ligands contained in the ion cluster increases. Usually, only a few ligands ($3\text{--}5$) are needed to make the agreement better than 1 kcal mol^{-1} .

Assuming that HSO_3^- is indeed observed as a ligand, only lower limits to $-\Delta G$ can be inferred from the stratospheric ion composition data, as equilibrium abundances may not be reached and the concentration of HSO_3^- is not easily derived. These lower limits (Table 2) are comparable to the corresponding values for H_2SO_4 -clustering, which suggests that the HSO_3^- -radical bonds strongly to HSO_4^- clusters. In turn, this implies that the heat of vaporization of HSO_3^- from an $\text{H}_2\text{SO}_4\text{--HSO}_3^-$ mixture should be $>18 \text{ kcal mol}^{-1}$.

Now we consider H_2O -clustering to negative ions. Measured hydrate-ion distributions are given in Table 2 and corresponding ΔG and ΔH values are listed in Table 3. The extent of hydration first seems to decrease as the cluster size increases and that it increases for $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_b$ clusters when $b \geq 2$. The former behaviour is known from laboratory studies^{20,24} of negative ion hydrates containing H_2O -ligands only. Unfortunately, no studies of mixed clusters containing H_2O - and H_2SO_4 -ligands are available. Qualitatively, a decrease of H_2O -bond energies with increasing cluster size should also be expected from a simple charge-dipole interaction model.

Evidently, if the observed increase of ΔH for $b > 2$ is real, it reflects the occurrence of strong cooperative bonding between H_2O - and H_2SO_4 -ligands. When compared with the heat of vaporization for bulk phase H_2O ($10.4 \text{ kcal mol}^{-1}$), our inferred ΔH -values for $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_3(\text{H}_2\text{O})_d$ ($d = 1, 2$) are markedly larger. Qualitatively, this may be expected from bulk phase considerations, as the $\text{H}_2\text{SO}_4\text{--H}_2\text{O}$ system has a large heat of mixing.

Our inferred values for ΔH of $13\text{--}14 \text{ kcal mol}^{-1}$ are much closer to the heat of vaporization of water from a bulk phase solution of $75\% \text{ H}_2\text{SO}_4$ and $25\% \text{ H}_2\text{O}$ (13 kcal mol^{-1}), which is thought to be the composition of the stratospheric aerosol. Additional strong support for the above interpretation comes from recent studies of negative ion hydration in our laboratory¹³ where it was found that ΔG values for the hydration of the ion series $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_b$, in fact, increase as b becomes >2 .

However, due to their small bond energies, negative ion hydrates may undergo substantial collisional dissociation while being sampled into the mass spectrometer. This is supported by simulation experiments¹³ in which it was found that in conditions similar to those for flight electric field-induced detachment of the weakly bonded H_2O molecules can be important. But this effect should at most raise the $-\Delta G$ values by 1 kcal mol^{-1} (90% dissociation) and otherwise not significantly affect our interpretations.

We now turn to HNO_3 -clustering. Here direct comparison with laboratory ΔH -data^{20,25} can be made, at least for $\text{NO}_3^-(\text{HNO}_3)_2$ and $\text{NO}_3^-(\text{HNO}_3)_3$ and good agreement is found. This supports our approach and suggests that no significant dissociation of these cluster ions occurs at the altitudes considered. This does not conflict with the above remarks on H_2O dissociation because HNO_3 bonds more strongly to negative ions than does H_2O . This conclusion is also supported by the observed large abundance of $\text{H}^+(\text{H}_2\text{O})_4$ ions (bond energy:

17 kcal mol⁻¹) which excludes significant dissociation of this ion. Thus, for ligands with bond strengths >17 kcal mol⁻¹ no significant dissociation seems to occur.

Again, in the case of HNO₃ an increase in ΔG seems to occur at large cluster sizes (HSO₄⁻(H₂SO₄)₃HNO₃). This may suggest that the heat of vaporization of HNO₃ from bulk phase H₂SO₄-HNO₃ is much higher than that for bulk phase HNO₃ which is only ~8 kcal mol⁻¹. It seems, therefore, that the observed heavy HSO₄⁻-clusters exhibit properties which are similar to those of the bulk phase.

Implications for aerosols

Our *in situ* composition measurements of stratospheric negative ions have important implications for stratospheric aerosols, concerning both their nature and their formation mechanism. Because thermodynamic properties of large ion clusters seem to approach those of the corresponding bulk phase, information on the composition of aerosols may be obtained from studies of large ion clusters existing in the same environment. The ion clusters studied in the present *in situ* measurements, however, are not very large and, therefore, their properties may still depart from those of the bulk phase. Thus, the following discussion should be treated with caution.

The largest negative ion clusters measured contain mostly H₂SO₄ and H₂O, and thus, they contain exactly the same molecular species which are thought to make up most of the stratospheric aerosol mass. However, the ligand shell of the ion cluster HSO₄⁻(H₂SO₄)₃(H₂O), for example, contains much less water (up to about 8% by mass) than the aerosol solution droplets (about 25%). This discrepancy may be partly due to cluster ion fragmentation occurring during the measurement, but most of the discrepancy appears real.

The detection of HSO₃ and the discovery that it bonds strongly to the large negative ion clusters suggests that the as yet unknown heat of vaporization of the HSO₃ radical is large, possibly comparable with that of H₂SO₄, which would imply that HSO₃ also undergoes efficient attachment to stratospheric aerosols.

If HSO₃ is, indeed, a precursor of stratospheric H₂SO₄, this may imply that HSO₃ attaches to aerosol solution droplets and is then converted to H₂SO₄ by chemical reactions in the liquid phase. Subsequently, H₂SO₄ may evaporate from the droplets and a condensation-evaporation equilibrium should be established according to temperature²⁹.

Thus, stratospheric aerosols may grow mostly by condensation of the HSO₃ radical and its hydrates rather than by H₂SO₄-condensation. The possibility of HSO₃-nucleation and -condensation was originally discussed by Friend *et al.*²⁶: their "radical nucleation" hypothesis being based on their laboratory studies of the SO₂-photooxidation.

In view of the large estimated aerosol surface area density at heights below ~30 km gas phase HSO₃-radicals may therefore preferably attach to aerosols rather than attaching to ions or reacting with OH.

At 20 km, for example, the estimated HSO₃-lifetimes against these processes are about 2 × 10⁴ s, 7 × 10⁴ s and 10⁵ s. At higher altitudes where the aerosol abundance decreases and where OH becomes more abundant, gas phase HSO₃ may preferably react with OH.

Another implication of the present ion composition measurements concerns polyion nucleation. It has been proposed¹¹ that the coagulation of large positive and negative ion clusters may lead to polyions or multi-ion complexes, which, in turn, may grow to the size of condensation nuclei. The first step in polyion-formation is the formation of a stable ion pair by recombination of a negative and a positive cluster ion. To prevent spontaneous neutralization and thereby destruction of the ion pair, the following criterion has to be fulfilled:

$$S > IP - EA + E_C \quad (8)$$

Table 3 Free energy- and enthalpy differences for ion clustering as inferred from stratospheric ion composition data (32.9 km altitude)

Ion	Ligands	$-\Delta G^\circ$	$-\Delta H^\circ$	$-\Delta_i^\circ$
HSO ₄ ⁻ (H ₂ SO ₄) ₂ -H ₂ SO ₄	3	14	19.8	
HSO ₄ ⁻ (H ₂ SO ₄) ₃ -H ₂ SO ₄	4	13	18.5	
HSO ₄ ⁻ (H ₂ SO ₄) ₄ -H ₂ SO ₄	5	13	18.9	
HSO ₄ ⁻ (H ₂ SO ₄) ₂ -HSO ₃	3	>12	>18.0	
NO ₃ ⁻ HNO ₃ -H ₂ O	2	7.1	12.9	
HSO ₄ ⁻ -HNO ₃ -H ₂ O	2	7.2	13.0	
HSO ₄ ⁻ (H ₂ SO ₄) ₂ -H ₂ O	3	6.3	12.1	
HSO ₄ ⁻ (H ₂ SO ₄) ₃ -H ₂ O	4	8.0	13.8	
HSO ₄ ⁻ (H ₂ SO ₄) ₄ -H ₂ O	5	8.1	13.9	
HSO ₄ ⁻ (H ₂ SO ₄) ₃ H ₂ O-H ₂ O	5	7.2	13.0	
NO ₃ ⁻ HNO ₃ -HNO ₃	2	12.6	18.4	18.3
NO ₃ ⁻ (HNO ₃) ₂ -HNO ₃	3	10.2	16.0	16.1
HSO ₄ ⁻ HNO ₃ -HNO ₃	2	11.2	17.0	
HSO ₄ ⁻ (H ₂ SO ₄) ₂ -HNO ₃	3	10.3	15.9	
HSO ₄ ⁻ (H ₂ SO ₄) ₃ -HNO ₃	4	11.1	16.9	

Values are given in units of kcal mol⁻¹. Laboratory data, if available, are also given for comparison. For the derivation of ΔH as common ΔS of 25 cal mol⁻¹ K⁻¹ was assumed.

where *IP* and *EA* are the ionization potential of the precursor neutral of the positive ion core and the electron affinity of the precursor neutral of the negative ion core. *E_C* is the energy released on formation of a chemical bond, and *S* is the total solvation energy or, in other words, the total energy for all bonds between ligand molecules and the core ions.

For HSO₄⁻(H₂SO₄)_b, H⁺(H₂O)_n and H⁺(CH₃CN)_l(H₂O)_m ions, the most prominent ions around 32–34 km, the inequality (8) becomes

$$S > 13.6 \text{ eV} \quad (9)$$

assuming that H⁺(CH₃CN)_l(H₂O)_m solvation energies are similar to those of H⁺(H₂O)_n. This is supported by recent laboratory measurements²⁷ of the competitive solvation of a proton by CH₃CN and H₂O. Here, a lower limit to *EA*(HSO₄) = 4.5 eV as measured by Viggiano *et al.*⁵ is assumed. *E₀* has been set equal to the H-HSO₄ bond strength.¹⁹ While the total solvation energy for H⁺(H₂O)_n is known, that for HSO₄⁻(H₂SO₄)_b ions is not known for *b* equal to 1, 2, and 3. Assuming $\Delta H = 1, 2, 3$ to be the same as that deduced from the present data for *b* = 4, 5 (~20 kcal mol⁻¹) upper limits to the critical sizes of cluster ions having the potential to form stable ion pairs can be estimated. The resulting minimum (*n*, *b*) pairs for critical clusters are: (2, 6), (3, 5), (4, 4), (5, 3) and (7, 2). Because ΔH values for HSO₄⁻(H₂SO₄)_b (*b* = 1, 2, 3) were underestimated and a lower limit to the electron affinity of HSO₄⁻ was used, the actual critical values may be smaller by about one ligand molecule.

In view of the fact that H⁺(H₂O)₄ and H⁺(H₂O)₅ are the most abundant proton hydrates around 32–34 km, it seems that ion-ion recombination of the newly observed large HSO₄⁻-clusters may mainly lead to stable ion pair formation. Taking fractional ion abundances equal to measured fractional count rates, the fraction of ion-ion recombinations able to form stable ion-pairs is estimated to be about 0.1–0.7 around 32–34 km.

Thus, at the altitudes considered, the abundant negative ion clusters seem to grow just large enough to allow for stable ion-pair formation on mutual recombination with the dominant positive ion clusters.

At still lower altitudes, NO₃⁻(HNO₃)_a and positive ion clusters may again ultimately become large enough to allow for high polyion formation rates. However, in this region, small polyions, due to the large aerosol surface area density, may rapidly attach to aerosols before they reach nuclei sizes.

Because negative ion clusters are largest around 30–35 km and because the size of positive ion clusters decreases as height increases the region around 30 km seems to be particularly favourable for polyion formation. Around this altitude, there may exist a layer of seed particles, which on mixing into the region below 30 km, where HSO_3 or H_2SO_4 become highly supersaturated, may initiate nuclei formation.

Conclusions

We have shown that *in situ* composition measurements of stratospheric ions can provide valuable information on thermodynamic properties of large ion clusters and aerosols. In addition, they yield new information on condensable vapours like H_2SO_4 and possibly also HSO_3 . Thus, *in situ* ion composition measurements strongly contribute to our understanding of

stratospheric trace gas and aerosol processes.

Substantial progress in the new field of stratospheric ion composition measurements can be expected when it becomes possible to lower ion detection limits and to reduce collisional dissociation occurring during ion sampling. Thus, it may be possible to detect even larger ion clusters having low abundances and bond energies.

We thank R. Turco, J. Friend, P. Crutzen and D. Ehhalt for helpful discussions. We also thank the technical staff of the Cosmophysics Department of the Max-Planck-Institute for Nuclear Physics, particularly K. Oberfrank, W. Thron, D. Dorflinger, H. Sauer, B. Preissler and K. Bechberger for their encouragement without which this project would not have been possible. Part of the project was funded by the Bundesministerium für Forschung und Technologie through Gesellschaft für Strahlen- und Umweltforschung.

Received 29 December 1981; accepted 31 March 1982.

1. Arnold, F. & Henschen, G. *Nature* **257**, 521 (1978).
2. Arnold, F., Fabian, R., Ferguson, E. E. & Joos, W. *Planet. Space Sci.* **29**, 195 (1981).
3. Turco, R. P., Hamill, P., Toon, O. B., Whitten, R. C. & Kiang, C. S. *J. atmos. Sci.* **86**, 699 (1979).
4. Viggiano, A. A. & Arnold, F. *Planet. Space Sci.* **29**, 895 (1981).
5. Viggiano, A. A., Perry, R. A., Albritton, D. L., Ferguson, E. E. & Fehsenfeld, F. C. *J. geophys. Res.* **85**, 4551 (1980).
6. Arnold, F. & Fabian, R. *Nature* **283**, 55 (1980).
7. McCrumb, J. L. & Arnold, F. *Nature* **294**, 136 (1981).
8. Arnold, F., Fabian, R. & Joos, W. *Geophys. Res. Lett.* **8**, 293 (1981).
9. Viggiano, A. A. & Arnold, F. *Geophys. Res. Lett.* **8**, 583 (1981).
10. Arijs, E., Nevejans, D., Frederick, P. & Ingels, J. *J. geophys. Res. Lett.* **8**, 121 (1981).
11. Arnold, F. *Nature* **284**, 610 (1980).
12. Arnold, F., Henschen, G. & Ferguson, E. E. *Planet. Space Sci.* **29**, 185 (1981).
13. Glebe, W. & Arnold, F. *Proc. Spring Meet. Arbeitsgemeinschaft für Extraterrestrische Forschung*, Köln (in the press).
14. Castleman, A. W. Jr, Davis, R. E., Munkelwitz, H. R., Tang, I. N. & Wood, W. P. *Int. J. chem. Kinet. Symp.* **1**, 629 (1975).
15. Moortgat, G. K. & Junge, C. E. *Pure appl. Geophys.* **115**, 759 (1977).
16. Rosen, J. M. & Hofmann, D. J. *J. geophys. Res.* **86**, 7399 (1981).
17. Rowe, B. R., Viggiano, A. A., Fehsenfeld, F. C., Fahey, D. W. & Ferguson, E. E. *J. Chem. Phys.* **76**, 742 (1982).
18. Stull, D. R. & Prophet, H. (eds), *JANAF Thermodynamical Tables* (Nat. Stand. Ref. Data Ser., National Bureau of Standards, Washington DC, 1971).
19. Bensen, W. W. *Chem. Rev.* **78**, 23 (1978).
20. Lee, N., Keesee, R. G. & Castleman, A. W. Jr *J. chem. Phys.* **72**, 1089 (1980).
21. Ayers, C. P., Gillett, R. W. & Gras, J. L. *Geophys. Res. Lett.* **7**, 433 (1980).
22. Kebarle, P. A. *Rev. phys. Chem.* **28**, 445 (1977).
23. Keesee, R. G., Lee, N. & Castleman, A. W. Jr *J. chem. Phys.* **13**, 2195 (1980).
24. Keesee, R. G., Lee, N. & Castleman, A. W. Jr *J. Am. chem. Soc.* **101**, 2559 (1979).
25. Davidson, J. A., Fehsenfeld, F. C. & Howard, C. J. *Int. J. Mass Spectrom. Ion Phys.* **9**, 17 (1977).
26. Friend, J. P., Barnes, R. A. & Vista, R. M. *J. phys. Chem.* **84**, 2423 (1980).
27. Böhringer, H. & Arnold, F. *Nature* **290**, 321 (1981).
28. Hamill, P., Turco, R. P., Toon, O. B., Kiang, C. S. & Whitten, R. C. *J. atmos. Sci.* (in the press).

A computational model of binocular depth perception

J. E. W. Mayhew* & H. C. Longuet-Higgins†

* Department of Psychology, University of Sheffield, Sheffield S10 2TN, UK
† Laboratory of Experimental Psychology, University of Sussex, Brighton, Sussex BN1 9QG, UK

We point out that the horizontal disparities between a pair of retinal images are inadequate for computing the three-dimensional structure of a scene unless supplemented by independent information about the distance and direction of the fixation point. We suggest that this supplementary information is derived not from non-visual sources, but from the vertical disparities of a few non-meridional image points. This hypothesis is shown to account quantitatively for Ogle's induced effect—the marked distortion of a scene by a vertically magnifying lens placed in front of one eye.

THE binocular perception of depth may be regarded as proceeding in two stages: the establishment of a point-by-point correspondence between the left and right retinal images to extract disparity information, and the interpretation of the disparities to yield a three-dimensional percept. The position of an image point on the left retina can be represented by its plane projective coordinates (x, y) ; if (x', y') are the coordinates of the corresponding image point on the right retina, then the difference $x' - x$ is described as the horizontal disparity between the two image points, and $y' - y$ as the vertical disparity. The 'correspondence problem'—that of computing the disparities—has been fully discussed in the literature^{1–5} and we shall not consider it further. Our present concern is the 'interpretation problem'—that of using the disparities for computing the three-dimensional structure of the scene.

As vertical disparities produced by local depth variations (in contrast to those produced by asymmetrical convergence) are commonly much smaller than horizontal ones^{1,8}—and

necessarily vanish on the horizontal meridian—it is often supposed that only the horizontal disparities are relevant to the binocular perception of depth. There are two difficulties with this hypothesis. First, the horizontal disparities supply insufficient information, even in principle, for computing the absolute distances and directions of a set of visible points; they must be supplemented by independent information about the distance and direction of the fixation point, but non-visual estimates of these 'viewing parameters' are notoriously unreliable. There is, however, a more serious difficulty: the binocular perception of a three-dimensional scene is profoundly altered by distorting one of the retinal images in the vertical dimension. If a cylindrical lens which induces a small vertical magnification is placed in front of one eye, then a visually textured surface in the fronto-parallel plane will acquire a pronounced tilt, as if it had been rotated about a vertical axis, although the horizontal disparities and the non-visual cues to fixation distance and direction are quite unaffected by the introduction of the lens.

The question therefore arises: could the vertical disparities supply the information about the viewing parameters that is required for deriving the structure of the scene from the horizontal disparities? The answer, as we shall see, is in the affirmative; and the hypothesis that the viewing parameters are derived from the vertical disparities rather than from non-visual information sources accounts quantitatively for the above phenomenon, which Ogle has described as the 'induced effect'¹.

It has recently been pointed out^{6,7} that a binocular system which takes account of the vertical as well as the horizontal image coordinates can in principle solve the interpretation problem without recourse to extra-retinal sources of information; and also⁸ that there exists a remarkably simple approximate method of deriving the position and orientation of a visually textured plane from the horizontal and vertical disparities of a small number of texture elements lying in the plane. Figure 1 illustrates the geometry of this situation. O and O' are the nodal points of the two eyes, and the observer is looking—not quite straight ahead—at the plane $Z = PX + QY + R$, where the Z axis joins the mid-point of O O' to the fixation point (0, 0, R) and makes an angle g with the forward direction. The capital letters refer to a coordinate system with the origin at the midpoint of the line joining the optical centres of the eyes. If the fixation distance R is large enough compared with the interocular distance I , then one can neglect, in the disparities, terms of the second order in I/R . It then becomes possible to use the algebra that Longuet-Higgins and Prazdny¹⁴ used in their analysis of the visual perception of motion; the problem being isomorphous with that presented by the stereoscopic viewing of an object at a moderate distance. If $\sin(g) = s$ and $\cos(g) = c$, one obtains, in this way,

$$H(x, y) = x' - x = [(Pc + s)x + Qcy + (c - Ps)x^2 - Qsxy]I/R \quad (1)$$

$$V(x, y) = y' - y = [sy + (c - Ps)xy - Qsy^2]I/R \quad (2)$$

These equations simplify if g is small enough for s and c to be replaced by g and 1 respectively, and for Pg and Qg to be neglected; in these conditions

$$H(x, y) = x' - x = [(P + g)x + Qy + x^2]I/R \quad (3)$$

$$V(x, y) = y' - y = (gy + xy)I/R \quad (4)$$

These equations immediately suggest a way of computing the unknown parameters P , Q , R and g . The equation for V implies that a plot of V/y against x gives a line of slope I/R and intercept gI/R , so that the values of R and g can be obtained from the vertical disparities of just two non-meridional points with different x coordinates (that is, points with $y \neq 0$ and $x_1 \neq x_2$). Explicitly:

$$R = I/V_{xy} \quad g = V_y/V_{xy} \quad (5)$$

where V_y and V_{xy} denote the coefficients of y and xy in the expression for V . Likewise, if H_x and H_y denote the coefficients of x and y in H , it follows from equation (3) that

$$P + g = H_x R/I, \quad Q = H_y R/I \quad (6)$$

Substituting from equation (5) into equation (6), we infer that

$$P = (H_x - V_y)/V_{xy}, \quad Q = H_y/V_{xy} \quad (7)$$

In short, the viewing parameters R and g can be directly computed from the vertical disparities alone; once they are known the geometrical parameters P and Q may be computed from either equation (6) or equation (7), in which the viewing parameters do not appear explicitly. A computer simulation implementing equations (5) and (7) has shown that the small-angle approximations underlying these equations are indeed adequate for a wide range of reasonable surface orientations, viewing distances and angles of gaze.

When a vertically magnifying lens is placed in front of one eye, the two retinal images become mutually incompatible, in the sense of ceasing to admit of a consistent three-dimensional interpretation. It is therefore quite surprising that the observer should obtain a clear three-dimensional percept, albeit a distorted one.

The induced effect therefore offers a challenge to theories of stereopsis. Would a 'seeing machine' which was based on a particular theory be subject to the induced effect, or would it arrive at a different interpretation of the two images, or fail to reach any interpretation at all? For example, machines based on the theories described in refs 6 and 7 would not be subject to the induced effect and therefore may be excluded as models of the human visual process. Similarly, a full implementation of equations (1) and (2) would not in general show the induced effect, whereas a machine implementing the approximate solutions [equations (3) and (4)] apparently interprets the two images in a similar way to that reported by human subjects. We shall now show that a system which computes P and Q directly from equation (7) will be subject to an induced effect of exactly the same nature and magnitude as that investigated by Ogle.

Suppose that a lens of horizontal magnification λ and vertical magnification μ is placed over the right eye. Then the new retinal coordinates for that eye will be

$$\hat{x} = \lambda x', \quad \hat{y} = \mu y' \quad (8)$$

so that the horizontal and vertical disparities will be

$$\hat{H} = \hat{x} - x = \lambda[(P + g)x + Qy + x^2]I/R + (\lambda - 1)x \quad (9)$$

$$\hat{Y} = \hat{y} - y = \mu(gy + xy)I/R + (\mu - 1)y \quad (10)$$

The new values of V_x and V_{xy} will therefore be

$$\hat{V}_x = \mu gI/R + (\mu - 1) = \mu(V_y + 1) - 1 \quad (11)$$

$$\hat{V}_{xy} = \mu I/R = \mu V_{xy} \quad (12)$$

and the new values of H_x and H_y will be

$$\hat{H}_x = \lambda(P + g)I/R + (\lambda - 1) = \lambda(H_x + 1) - 1 \quad (13)$$

$$\hat{H}_y = \lambda QI/R = \lambda H_y \quad (14)$$

The apparent values of P , Q and R will therefore be

$$\hat{P} = (\hat{H}_x - \hat{V}_x)/\hat{V}_{xy} = [\lambda(H_x + 1) - \mu(V_y + 1)]/\mu V_{xy} \quad (15)$$

$$\hat{Q} = \hat{H}_y/\hat{V}_{xy} = \lambda H_y/\mu V_{xy} = \lambda Q/\mu \quad (16)$$

$$\hat{R} = I/\hat{V}_{xy} = I/\mu V_{xy} = R/\mu \quad (17)$$

If λ and μ are both ≈ 1 , the values of Q and R are not much affected, but with P it is an entirely different matter. To see

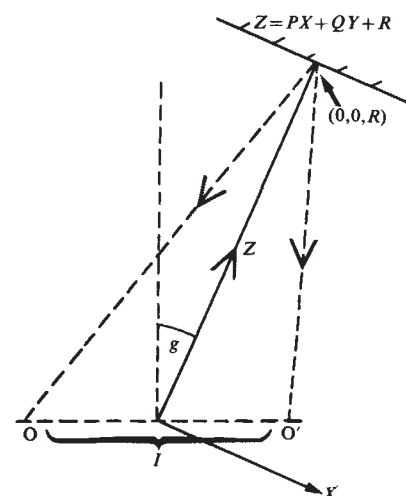


Fig. 1 The coordinate system has origin at the midpoint of the line joining the optical centres of the eyes (O and O'). The line joining this point to the point of fixation (0, 0, R) defines the Z axis; the X axis lies in the plane containing the line OO' and fixation point; and the line normal to this plane is the Y axis. The angle of gaze g , is the angle OO' makes with the X axis. The analysis assumes that the eyes are free to rotate around their X and Y axes but not their Z axes.

this, suppose that P is actually zero, as it will be for a fronto-parallel plane. Then, according to equation (7), H_x and V_y must be equal, and so the apparent value of P , given by equation (15), will be

$$\hat{P} = (\lambda - \mu)(1 - V_y)/\mu V_{xy} \quad (18)$$

As V_y is of order I/R , and λ and μ are close to 1, we may neglect V_y in the numerator and infer that

$$\hat{P} \approx (\lambda - \mu)/V_{xy} = (\lambda - \mu)R/I \quad (19)$$

Equation (19), with μ set equal to 1, becomes Ogle's own expression for the apparent tilt produced by a horizontally magnifying lens—an effect which can be understood in purely geometrical terms and which he therefore named the 'geometric effect'. Various predictions follow, and all are confirmed by Ogle's observations:

- (i) The apparent tilt is proportional to the difference between the horizontal and vertical magnification, so that for a given value of R a vertical magnification induces an apparent tilt equal and opposite to that produced by an equivalent horizontal magnification. (The induced effect actually falls off at high magnification; this is not predicted by equation (19), but is presumably connected with the fact that $\hat{V}_y/\hat{V}_{xy}$ then implies an implausibly large value for the angle of gaze.)
- (ii) An ordinary spherical lens, with $\lambda = \mu$, induces no tilt.
- (iii) When $\mu = 1$ and λ near 1 the induced tilt is proportional both to $\lambda - 1$ and to R . At a viewing distance $R = 40$ cm, Ogle found that the angle of tilt increased by 3–3.5° for every 1% increase in magnification. With an interocular distance $I = 6.5$ cm, equation (19) predicts a value for this ratio of

$$(180/\pi) \times (40/6.5) \times (1/100) = 3.53^\circ \text{ per \% magnification}$$

In his experiments, Ogle actually used a null method of estimating the induced tilt of a vertical plane: the subject had to tilt the plane himself (in the opposite direction) until the apparent slope P became zero, that is, until $\hat{H}_x = \hat{V}_y$. This relationship entails that

$$\lambda(H_x + 1) = \mu(V_y + 1) \quad (20)$$

or

$$(\lambda - \mu) = \lambda H_x - \mu V_y = [\lambda(P + g) - \mu g]I/R \quad (21)$$

Remembering that $\lambda = 1$ and $\mu = 1$, we get

$$\mu - \lambda \approx PI/R \quad (22)$$

showing that the actual value of P required to make $P = 0$ is

$$P = (\mu - \lambda)R/I \quad (23)$$

This is, indeed, equal and opposite to the apparent tilt of a fronto-parallel plane under identical viewing conditions.

Discussion

It has commonly been believed (see ref. 9 for a review) that only the horizontal disparities convey useful information about the distances of objects, and only qualitative information at that. Vertical disparities, which arise when an object is nearer one eye than the other, have been regarded merely as something that the visual system must be capable of allowing for in fusing the two images into a unitary percept. Our approach to the problem, from an artificial intelligence standpoint, has led to a computational solution of the interpretation problem which accounts quantitatively for the induced effect. Various explana-

tions of this effect have been proposed, both by Ogle and by others^{1,2,9–12}; but none, we feel, follows so naturally from a theory of normal stereoscopic depth perception as the explanation proposed above (the limitations of a recent theory¹² which attempts to explain the induced effect without using vertical disparities have been criticised elsewhere¹¹). Ogle's explanation, like ours, agrees quantitatively with the psychophysical magnitude of the induced effect, but requires the use of the vertical disparities on the vertical meridian to compute the compensatory rotation of the Veith-Müller circle necessary to maintain the correct egocentric localization (this corresponds to the direction of gaze in our expressions). However, Ogle's theory fails to exploit the vertical disparity information at other retinal locations and thus does not solve for the other viewing parameter, distance, which remains as an unknown scaling variable in his expression. Note that the retinal eccentricity of the location at which the disparity information is measured is of fundamental importance in our expressions relating horizontal and vertical disparity to the viewing and scene parameters [equations (1)–(4)]. Without the terms derived solely from the retinal eccentricity, the system of equations would be homogeneous and without a unique solution. The emphasis on a coordinate system derived from concern with corresponding points—the horopter and the Veith-Müller circle—may have masked the importance of these terms from Ogle's consideration, and may explain why, when he was so close to a complete solution, he failed to find it.

As our theory is formally complete, it predicts that the induced effect will result from any stimulus which supplies misleading cues about the viewing parameters—the distance and direction of the fixation point. Although the vertical disparities from any two retinal locations could be used to solve for the viewing parameters, and thus the computation could be essentially local to a particular region of retina, the results of the computation have global implications which can provide powerful constraints and checks on its consistency. In contrast to the interpretation of horizontal disparities where the results in one region of the retina exert little constraint on the interpretation at another region, the interpretation of the vertical disparities should give the same result everywhere. There are obvious advantages in making viewing system parameters globally available, and so one might expect that the induced effect would act on the whole visual field even though vertical disparity information could only be derived from a particular region of it. Ogle¹ reports an experiment confirming this expectation. Vertical disparities from one part of the visual field changed the apparent depths of vertical rods (which of course cannot provide vertical disparity information) in accord with the predictions of the induced effect. If the stimulus situation presents incompatible vertical disparity information at different parts of the visual field or, alternatively, vertical disparity information whose interpretation results in grossly implausible viewing system parameters, one might expect some difficulty in perceiving the induced effect. This may explain why there may be some difficulty in producing the induced effect without the aid of a cylindrical lens^{12,13}.

Finally, it is of interest that Ogle reports excellent agreement in estimates of stereoacuity derived from experiments investigating the induced effect and other psychophysical procedures, which correspond to a sensitivity of about 0.2% magnification between the two eyes' images. The constraints that this degree of resolution imposes on the implementation of the processes described above are currently the subject of computational and psychophysical experimentation.

Received 13 October 1981; accepted 26 March 1982.

1. Ogle, K. N. *Binocular Vision* (Saunders, Philadelphia, 1950).
2. Julesz, R. *Foundations of Cyclopean Perception* (University of Chicago Press, 1971).
3. Nelson, J. J. *theor. Biol.* **49**, 1–88 (1975).
4. Marr, D. & Poggio, T. *Proc. R. Soc. R204*, 301–328 (1979).
5. Mayhew, J. E. W. & Frisby, J. P. *Art. Intelligence* **23**.
6. Longuet-Higgins, H. C. *Nature* **293**, 133–134 (1981).

7. Longuet-Higgins, H. C. *Perception* (in the press).
8. Mayhew, J. E. W. *Perception* (in the press).
9. Foley, J. M. *Psychol. Rev.* **87**, 411–434 (1980).
10. Nelson, J. J. *theor. Biol.* **66**, 203–266 (1977).
11. Mayhew, J. E. W. & Frisby, J. P. *Vision Res.* (in the press).
12. Arditi, A., Kaufman, L. & Movshon, J. A. *Vision Res.* **21**, 755–764 (1981).
13. Westheimer, G. *Invest. Ophthalm. Visual Sci.* **17**, 545–551 (1978).
14. Longuet-Higgins, H. C. & Prazdny, K. *Proc. R. Soc. B208*, 385–397 (1980).

LETTERS TO NATURE

Could primordial black holes be the source of the cosmic ray antiprotons?

Michael S. Turner

Astronomy and Astrophysics Center, The University of Chicago, Chicago, Illinois 60637, USA

Antiprotons ($\bar{p}$) are expected to be found in the cosmic rays at the level of $\sim 10^{-4}$ relative to protons (p) as they should be produced by cosmic-ray proton collisions with the interstellar medium¹⁻³. The recent observations of Buffington *et al.*⁴ are puzzling, however, because the total $\bar{p}$ flux observed is ~ 10 times larger than predicted, and the shape of the $\bar{p}$ spectrum differs greatly from that predicted. We explore here the possibility that primordial black holes (PBHs) of mass $\sim 10^{13}$ – 10^{15} g which evaporated after decoupling ($t \gtrsim 10^{14}$ s) are the primary source of the cosmic ray $\bar{p}$ s. The PBH scenario predicts a universal $\bar{p}/p$ ratio of $\sim 10^{-4}$ – 10^{-8} .

Cosmic-ray protons are generally believed to traverse ~ 5 g cm⁻² during their life in the Galaxy. Cosmic-ray proton collisions with the interstellar gas are therefore predicted to produce an $\bar{p}$ flux of $\bar{p}/p \approx 2 \times 10^{-4}$ for $\bar{p}$ kinetic energies above a few GeVs. The predicted $\bar{p}/p$ ratio falls very steeply for $\bar{p}$ energies below about 1 GeV (refs 1–3). The recent measurements of Buffington *et al.*² in the energy range 130–320 MeV indicate a $\bar{p}$ flux of $\sim 2 \times 10^{-5}$ cm⁻² sr⁻¹ s⁻¹ GeV⁻¹, which is several orders of magnitude larger than the flux expected if the $\bar{p}$ s are cosmic ray secondaries. The three measurements made so far span the energy range from ~ 100 MeV to ~ 10 GeV, and are consistent with a constant $\bar{p}/p$ ratio of $\sim 3 \times 10^{-4}$. The integrated $\bar{p}$ flux is about a factor of 10 larger than is predicted.

There are three possible explanations of the discrepancy. (1) The measurements are wrong. The results for the energy range of a few to 10 GeV are only a factor of two to three higher than the predicted flux^{5,6}. Buffington *et al.*'s⁴ measurement shows greater divergence. (2) The $\bar{p}$ s observed are indeed secondary particles, but the standard picture of cosmic ray propagation must be modified. For example, some of the cosmic ray protons would have to traverse significantly more grammage than ~ 5 g cm⁻² thus producing more $\bar{p}$ s⁷⁻⁹, and a stochastic mechanism¹¹ to decelerate the $\bar{p}$ s would be required to explain the high flux of low-energy $\bar{p}$ s Buffington *et al.*⁴ observe. (3) The $\bar{p}$ s detected are primary particles originating from the big bang, or from regions of antimatter. If their origin is the big bang, then there is the problem of storing them so that they escape annihilation¹², and then subsequently injecting them into the Galaxy. If the observed $\bar{p}$ s originate from antigalaxies¹³, then it is difficult to understand why the ${}^4\bar{\text{He}}/{}^4\text{He}$ ratio is $< 2.2 \times 10^{-5}$ (95% confidence, ref. 4), rather than being equal to the $\bar{p}/p$ ratio $\approx 3 \times 10^{-4}$.

Here I discuss three models in which evaporating PBHs of mass $\sim 10^{13}$ – 10^{15} g radiate $\bar{p}$ s sufficiently late in the history of the Universe that the $\bar{p}$ s avoid annihilation, leading to a universal abundance of $\bar{p}/p \sim 10^{-4}$ – 10^{-8} . The initial spectrum of PBHs required to explain the observations is consistent with previous constraints on the spectrum of PBHs¹⁴, and could also lead to the production of a significant baryon asymmetry¹⁵ (much earlier in the history of the Universe, $t \sim 10^{-35}$ s). Because the PBHs must have a temperature of ≥ 1 GeV when they produce the $\bar{p}$ s, antinuclei will not be produced in the process, thus explaining the absence of antinuclei.

First, consider the question of producing $\bar{p}$ s that do annihilate with the ambient protons in the Universe. The $\bar{p}$ annihilation rate Γ_{ann} is given by

$$A6\Gamma_{\text{ann}} = n_p(\sigma v)_{\text{ann}} \quad (1)$$

The number density of protons, n_p , is just $(n_p/n_\gamma) \times n_\gamma \approx (3 \times 10^{-10}) \times (2\zeta(3)T^3/\pi^2)$, where $n_p/n_\gamma \approx (3-5) \times 10^{-10}$ (J. Yang *et al.*, in preparation) and the second factor is the number density of photons at temperature T ; $\zeta(3) \approx 1.202$. The annihilation cross-section times the relative velocity $(\sigma v)_{\text{ann}} \approx m_\pi^{-2}$. Annihilations will not occur at a significant rate if $\Gamma_{\text{ann}} < H \equiv \dot{R}/R = 1.66 g_*^{1/2} T^2/m_{\text{pl}}$ (the expansion rate). Here $\hbar = c = k_B = 1$, $m_{\text{pl}} = G^{-1/2} \approx 10^{19}$ GeV and g_* is the total number of effective degrees of freedom of all the relativistic species at a temperature T . The condition $\Gamma_{\text{ann}} < H$ requires

$$T \lesssim 10^{10} (\pi^2/2\zeta(3)) m_\pi^2/m_{\text{pl}} \sim 0.1 \text{ eV} \quad (2)$$

This temperature roughly corresponds to that of decoupling ($T \sim 1,000$ K), and to a time $t \sim 10^{14}$ s. Essentially all the $\bar{p}$ s produced after $t \sim 10^{14}$ s will avoid subsequent annihilation, while $\bar{p}$ s produced earlier will be annihilated. (If $n_p/n_\gamma \geq (3-5) \times 10^{-10}$, then the Universe will be matter-dominated when the annihilations cease and the analysis above must be modified. If, however, the Universe is nucleon-dominated now, then the time at which annihilations cease is independent of n_p/n_γ . If the Universe is not nucleon-dominated and $\Omega h^2 \approx 1$, $t \approx 3 \times 10^{12}$ s).

Hawking¹⁶ showed that a black hole of mass m should radiate like a blackbody with temperature

$$T \approx 0.1 m_{\text{pl}} (m_{\text{pl}}/m) \quad (3)$$

and evaporate in a time¹⁷

$$\tau \approx m_{\text{pl}}^{-1} (m/m_{\text{pl}})^3 (100/g_*) \quad (4)$$

where g_* is the number of species with mass less than T . Of course, this process is only important for black holes of mass $m \ll 1 M_\odot$. Such black holes must be primordial (produced from density perturbations in the early Universe; see, for example, ref. 13), as known astrophysical processes can only produce black holes of mass $M \gtrsim \text{few } M_\odot$.

Equation (4) shows that PBHs are ideal candidates for producing $\bar{p}$ s after $t \sim 10^{14}$ s, as they have a built-in clock. PBHs of mass $m \gtrsim 10^{18} m_{\text{pl}} \sim 10^{13}$ g will evaporate after $t \sim 10^{14}$ s. A PBH of mass $m \sim 10^{18} m_{\text{pl}}$ has a temperature of ~ 1 GeV. As it radiates, its mass decreases and its temperature increases. If we assume that for $T \gtrsim 1$ GeV about one-tenth of the particles the PBH radiates are $\bar{p}$ s¹⁴, and the typical $\bar{p}$ momentum is $\sim T$, then the spectrum of $\bar{p}$ s produced is

$$dn_{\bar{p}} \approx 2 \times 10^{36} p^{-3} dp \quad (5)$$

where the momentum p is measured in GeV. A PBH more massive than $10^{18} m_{\text{pl}}$ will not emit $\bar{p}$ s until its mass becomes less than $\sim 10^{18} m_{\text{pl}}$, and then it will emit the same spectrum as a PBH of that mass, given by equation (5).

Following Carr¹⁴, we parameterize the initial PBH spectrum as

$$\frac{dn(m)}{n_\gamma} = \zeta \left(\frac{m}{m_{\text{pl}}} \right)^{-\alpha} \frac{dm}{m_{\text{pl}}} \quad (6)$$

where $dn(m)/n_\gamma$ is the number of PBHs per photon in the mass interval $m \rightarrow m + dm$, ζ is the normalization ($\approx$ fraction of the mass density of the universe in PBHs at $t_{\text{pl}} \sim 10^{-43}$ s) and α is the index of the spectrum. The most stringent constraint on ζ and α results from the insistence that PBHs of mass $\sim 10^{15}$ g do not contribute excessively to the diffuse γ -ray background¹⁴,

$$\zeta \lesssim 10^{20\alpha-75} \quad (7)$$

Of course, ζ must also be < 1 . [In equations (6) and (7) it is assumed that the number of PBHs per photon has remained constant (neglecting evaporation). This relies on the assumption that the expansion of the Universe has been approximately isentropic, so that the number of photons in the Universe has not increased significantly.]

Summing the contribution of all PBHs more massive than $\sim 10^{13}$ g and taking into account the subsequent cosmological redshift of the $\bar{p}$ momenta, we find that, at present,

$$\frac{dn_{\bar{p}}}{n_{\gamma}} \approx 10^{55.5-19.5\alpha} \zeta KE^{-3/4-\alpha/4} dKE \quad (8)$$

for $\bar{p}$ kinetic energies (measured in GeV) in the range $1 \text{ GeV} \geq KE \geq 10^{-6} \text{ GeV}$. The lowest-energy $\bar{p}$ s are those that were emitted with momentum $\sim 1 \text{ GeV}$ by 10^{13} g PBHs just after decoupling; at present they have $KE \sim 10^{-6} \text{ GeV}$ ($v \sim 10^{-3}c$). For relativistic $\bar{p}$ s ($E \geq \text{few GeV}$),

$$\frac{dn_{\bar{p}}}{n_{\gamma}} \approx 10^{55.5-19.5\alpha} \zeta E^{-3} dE \quad (9)$$

None of the $\bar{p}$ s produced by PBH evaporations would have initially collapsed into galaxies with the baryons, both because the $\bar{p}$ s are emitted during the epoch of galaxy formation ($z \leq 10^3$) and since they are relativistic when they are emitted they are moving too fast to be captured in galaxies or protogalaxies. The entire spectrum, equations (8) and (9), is primarily due to the $\sim 10^{13}$ g PBHs which evaporated shortly after decoupling.

The universal abundance of $\bar{p}$ s relative to photons is obtained by integrating over the spectrum

$$\frac{n_{\bar{p}}}{n_{\gamma}} \approx 10^{54.5-18\alpha} \zeta \leq 10^{-13} \quad (10)$$

and it is dominated by the low-energy $\bar{p}$ s ($v \approx 10^{-3}c$). The upper limit in equation (10) is obtained by using the maximum value of ζ consistent with the limit derived from the diffuse γ -ray background ($\zeta = 1$, $\alpha = 3.75$). Equation (10) translates into a universal $\bar{p}/p$ ratio $\leq 10^{-4}$.

Antiprotons are not annihilating rapidly enough now to deplete their abundance significantly; however, the π^0 s produced in $\bar{p}p$ annihilations contribute to the diffuse γ -ray background when they decay. To compute the annihilation rate one must use the nonrelativistic annihilation cross-section, $(\sigma v)_{\text{ann}} \approx 7 \times 10^{-17} \text{ cm}^3 \text{ s}^{-1} (c/v)$ (ref. 18). The contribution to the γ -ray background today is

$$I \approx [(\bar{p}/p)/10^{-4}] 3 \times 10^{-5} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad (11)$$

The flux observed at $E_{\gamma} \sim 100 \text{ MeV}$ is about $3 \times 10^{-5} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$ (ref. 19). The γ -ray background thus places an upper bound on the universal $\bar{p}/p$ ratio of $\sim 10^{-4}$ ($v/10^{-3}c$), regardless of their source (assuming that they are non-relativistic; see also ref. 12).

From the abundance of $\bar{p}$ s due to PBH evaporations [equations (8) and (9)], one must calculate the resulting galactic cosmic ray $\bar{p}$ flux. For reference, the observed $\bar{p}$ flux corresponds to a local antiproton number density of

$$n_{\bar{p}|\text{galaxy}} \approx 10^{-14} \text{ cm}^{-3} \quad (12)$$

essentially all in the energy range $0.1\text{--}1 \text{ GeV}$. If the Galaxy is taken to be a sphere of radius $\sim 30 \text{ kpc}$, and the $\bar{p}$ number density constant, there are about 10^{55} $\bar{p}$ in the Galaxy today. The escape time for cosmic rays of energy greater than several GeV is $\sim 10^7$ yr; however, lower-energy cosmic rays may be trapped for longer. In any case, the annihilation time scale for an $\bar{p}$ in the Galaxy is $\sim 10^8$ yr. Thus a reasonable estimate for the lifetime of an $\bar{p}$ in the Galaxy is $10^7\text{--}10^8$ yr, and so $\bar{p}$ s of energy $0.1\text{--}1.0 \text{ GeV}$ must be replenished at the rate of

$$\frac{dN_{\bar{p}}}{dt} \approx 10^{39}\text{--}10^{40} \text{ s}^{-1} \quad (13)$$

This is the rate at which the ambient intergalactic $\bar{p}$ s must be injected into the Galaxy.

There are three possible injection models: (1) low-energy ($v \approx 10^{-3}c$) $\bar{p}$ s are captured by the Galaxy and then accelerated up to $\sim 0.1\text{--}1.0 \text{ GeV}$; (2) $\bar{p}$ s in the energy range $0.1\text{--}1.0 \text{ GeV}$ are captured by the Galaxy; (3) PBHs of mass $\sim 10^{15}$ g which

are in our Galaxy and are evaporating today supply the required flux of $\bar{p}$ s. All three models are viable, and have associated uncertainties and difficulties.

In the case of model (1), the total flux of low-energy $\bar{p}$ s ($v \approx 10^{-3}c$) on the Galaxy is given by

$$\frac{dN_{\bar{p}}}{dt} \approx 10^{111-18\alpha} \zeta \text{ s}^{-1} \approx 10^{44} \text{ s}^{-1} \quad (14)$$

To supply the required number of $\bar{p}$ s in the energy range $0.1\text{--}1.0 \text{ GeV}$, the efficiency of capture and acceleration must be $> 10^{-4}$.

In model (2), the total flux of $\bar{p}$ s in the energy range $0.1\text{--}1.0 \text{ GeV}$ on the Galaxy is given by

$$\frac{dN_{\bar{p}}}{dt} \approx 10^{115-19.5\alpha} \zeta \text{ s}^{-1} \approx 10^{43} \text{ s}^{-1} \quad (15)$$

In this case to supply the required numbers of $\bar{p}$ s in the range $0.1\text{--}1.0 \text{ GeV}$, the probability that a given $\bar{p}$ with energy $\sim 0.1\text{--}1.0 \text{ GeV}$ incident on the Galaxy enters the galactic magnetosphere must be $\geq 10^{-4}$. In both models (1) and (2) the $\bar{p}$ s must scatter off inhomogeneities in the galactic magnetic field in order to become galactic cosmic rays. This process is not well understood, and remains the biggest uncertainty in the estimates for models (1) and (2). In both models the $\bar{p}$ s come primarily from $\sim 10^{13}$ g PBHs which evaporated at $t \sim 10^{14}$ s.

In the case of model (3), although the bulk of the $\bar{p}$ s in the intergalactic medium are produced by $\sim 10^{13}$ g PBHs which evaporate just after decoupling, $\sim 10^{15}$ g PBHs which are evaporating now are the primary source of galactic $\bar{p}$ s. If we assume that these PBHs collapsed with the baryons during galaxy formation, then their local number density is enhanced by a factor of $\sim 10^6$ relative to equation (6). The $\bar{p}$ s such PBHs radiate now have energies $0(1 \text{ GeV})$, and are already inside the Galaxy. The total flux of $\bar{p}$ s from $\sim 10^{15}$ g PBHs in the Galaxy is

$$\frac{dN_{\bar{p}}}{dt} \sim 10^{117-20\alpha} \zeta \text{ s}^{-1} \approx 10^{42} \text{ s}^{-1} \quad (16)$$

For the maximum value of ζ , the total flux is $10^2\text{--}10^3$ larger than required. Thus one could improve Carr's constraint¹⁴ on ζ slightly, $\zeta \leq 10^{-77+20\alpha}$. This model requires the number density of 10^{15} g PBH to be within $\sim 10^2\text{--}10^3$ of the maximum allowed number; it has the advantage that the $\bar{p}$ s do not have to penetrate the galactic magnetosphere to become cosmic rays, hence avoiding the uncertainties involved with this process.

It is therefore possible for the evaporation of $\sim 10^{13}\text{--}10^{15}$ g PBHs to produce the observed, but unexplained, large flux of $\sim 0.1\text{--}1.0 \text{ GeV}$ $\bar{p}$ s. The number density of PBHs must be within a factor of $\sim 10^4$ of the upper limit allowed by the γ -ray background constraint¹⁴. The mass spectrum of PBHs can have an index α in the range $2\text{--}3.75$. In the first and second models α must be ~ 3.75 unless the injection of $\bar{p}$ s into the Galaxy is very efficient. If PBH evaporations are responsible for the cosmic ray $\bar{p}$ s, then there should be a universal $\bar{p}/p$ ratio of $\sim 10^{-4}\text{--}10^{-8}$ in the intergalactic medium. In two of the models discussed, intergalactic $\bar{p}$ s must penetrate the galactic magnetosphere to become cosmic rays. In the third model, the cosmic ray $\bar{p}$ s are produced by $\sim 10^{15}$ g PBHs within the Galaxy which are evaporating today. None of the PBH models is compelling, but they are plausible, and perhaps even intriguing.

After completing this paper, I realised that Kiraly *et al.*²⁰ have also suggested PBHs as a source of the cosmic ray antiprotons. The model they discuss is similar to my model, in which PBHs evaporating in the Galaxy today are responsible for the cosmic ray $\bar{p}$ s.

I thank Tom Gaisser for introducing me to this problem and for beneficial discussions, and Gary Steigman for useful comments. This work was supported in part by the DOE (ACO2-80ER10773) at University of Chicago and by NSF grant PHY 77-27084 at University of California, Santa Barbara.

Received 10 November 1981; accepted 17 February 1982.

1. Gaisser, T. K. & Levy, E. H. *Phys. Rev. D* **10**, 1731 (1974).
2. Stephens, S. A. *Astrophys. Space Sci.* **76**, 87 (1981).
3. Szabelski, J., Wdowczyk, J. & Wolfendale, A. W. *Nature* **285**, 386 (1980).
4. Buffington, A. & Schindler, S. M. *Astrophys. J. Lett.* **247**, L105 (1981).
5. Golden, R. L. *et al. Phys. Rev. Lett.* **43**, 1196 (1979).
6. Bogomolov, E. A. *et al. Proc. 16th int. Cosmic Ray Conf.*, Kyoto, Vol. 1, 33 (1979).
7. Rasmussen, I. L. & Peters, B. *Nature* **258**, 412 (1975).
8. Peters, B. & Westergaard, N. J. *Astrophys. Space Sci.* **48**, 21 (1977).
9. Stephens, S. A. *Astrophys. Space Sci.* **76**, 87 (1981).
10. Steigman, G. *Astrophys. J. Lett.* **217**, L131 (1977).
11. Szabelski, J., Wdowczyk, J. & Wolfendale, A. W. *Nature* **285**, 386 (1980).
12. Steigman, G. A. *Rev. Astr. Astrophys.* **14**, 339 (1976).
13. Stecker, F. W., Protheroe, R. & Kazanas, D. *Proc. 17th int. Cosmic Ray Conf.*, Paris (in the press).
14. Carr, B. J. *Astrophys. J.* **201**, 1 (1975); **206**, 8 (1976).
15. Turner, M. S. *Phys. Lett.* **89B**, 155 (1979).
16. Hawking, S. W. *Nature* **248**, 30 (1974).
17. Page, D. N. *Phys. Rev. D* **13**, 198 (1975).
18. Morgan, D. L. Jr & Hughes, V. W. *Phys. Rev. D* **2**, 1389 (1970).
19. Fitchel, C. E. *et al. Astrophys. J.* **198**, 163 (1975).
20. Kiraly, P., Szabelski, J., Wdowczyk, J. & Wolfendale, A. W. *Nature* **293**, 120 (1981).

Isotopically distinguishable carbon phases in the Allende meteorite

P. K. Swart, M. M. Grady & C. T. Pillinger

Planetary Sciences Unit, Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK

The stepwise oxidation of the Allende meteorite reported here has shown that there are at least three isotopically distinct carbon phases, each combustible over different temperature ranges. The temperature of combustion is dependent both on grain size of the fragments studied and/or on the nature of the carbonaceous material and its availability to the oxygen. Thus we attempt here to correlate the isotopic signatures recognized with phases whose existence has been postulated to explain noble gas systematics^{1,2} and to provide some insight into the nature and location of these phases. Most of the carbon ($\delta^{13}\text{C} = -20\%$), possibly a highly cross-linked polymer, exists predominantly in the matrix and has been reported by others³ to be the host of most of the normal planetary noble gases. Carbon protected in mineral grains may be isotopically heavier at -13 to -7% , but is still more easily oxidized than graphite when exposed by HF/HCl demineralization. A very easily oxidized phase, as yet unidentified, has a $\delta^{13}\text{C}$ value typically ~ -23 to -27% .

Carbon isotopic analyses of carbon-rich meteorites (CI, CM2, C3 and type 3 unequilibrated ordinary chondrites), apart from H_3PO_4 -released fractions⁴, have revealed a very uninspiring range of values. In fact, after removal of obviously separable minor components, such as carbonate and soluble organics, some of which may be contaminants, bulk carbon usually has an isotopic composition of between -5 and -28% and shows no coherent trend relative to meteorite group or petrological type⁵. The bulk of the carbon in chondrites is largely uncharacterized and only operationally defined. Thus, being insoluble in mineral acids and organic solvents, it can only be some allotrope or polymer of carbon. The root cause of the unrepresentative isotopic compositions may be that bulk carbon is an unresolved mixture in such meteorites. With material of several different isotopic compositions present in varying proportions, bulk values would be at best confused. Recognition of several carbonaceous noble gas host-phases within a single meteorite^{1,2} apparently endorses such an interpretation. Obviously a method capable of further resolving mixtures within the bulk carbon would possibly allow isotope signatures to be recognized.

The Allende C3V meteorite has a reported carbon isotope value of between -16.4 and -19.5% (refs 6–8). Much effort has been made recently to separate carbon-rich phases, thought to be among the hosts of anomalous noble gases, from the meteorite^{1,2}. In one such study, a separate (AMII)⁹ was combusted and carbon isotopic analyses of the CO_2 produced reported by Ott *et al.*². The data showed increasing heavy isotope

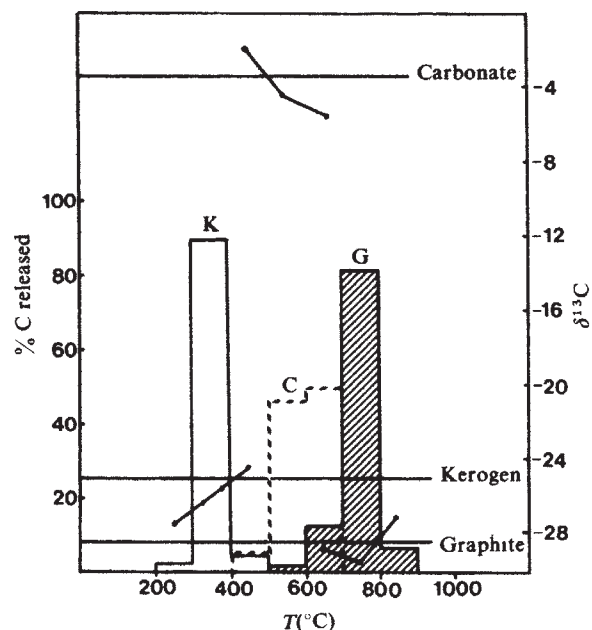


Fig. 1 Stepwise oxidation of three standard carbonaceous substances. Kerogen (type I kerogen extracted from Upper Jurassic Kimmeridge Clay); PSU-1 carbonate (sparry calcite standard); and graphite (NBS-21). The $\delta^{13}\text{C}$ values obtained by a one-step total extraction are shown as horizontal lines.

abundance with temperature. At the time, this change was thought to be a result of kinetic isotope effects and hence dismissed as of no significance.² Recently, however, we have obtained evidence which suggests that such an interpretation may have been premature. We performed a series of stepwise oxidations using various bulk sieve fractions of the Allende meteorite ($>50\mu\text{m}$, $<50\mu\text{m}$ and $<25\mu\text{m}$) as well as a sample from which silicate had been partially removed by HF/HCl treatment, in an effort to bridge the gap between fully demineralized and bulk samples. The method involved is similar to that used by others^{2,3} for stepwise oxidations and involves the combustion of each sample for 30 min in an atmosphere of O_2 (>20 mbar) derived from the decomposition of CuO . Gases evolved are frozen into a variable temperature cryogenic trap and excess O_2 adsorbed onto a CuO/Cu furnace. Carbon dioxide produced is separated from SO_2 and measured with a capacitance manometer. Isotopic measurements were performed using a VG Micromass 602-E and data reported in $\%$ relative to PDB. To evaluate the extent of kinetic isotope effects during stepwise combustion several carbonaceous substances were analysed using the same method.

Stepwise carbon release patterns and isotopic compositions for the standard and meteorite samples are shown in Figs 1 and 2. For the $<50\mu\text{m}$ fraction, carbon yield and isotopic values, calculated by summing the individual temperature steps, agree well with measurements made during conventional extraction procedures; bulk measurements have not been made on the other fractions. Unfortunately, the higher temperature extractions of the HF/HCl-treated sample were contaminated with fluorinated species and no isotopic data were obtained. However, on the basis of previous results², we assume that the 500 – 800°C fractions would have contained slightly heavier carbon. We discount the interpretation that an increase in heavy isotope abundance with temperature of oxidation, seen in all cases, merely reflects kinetic isotope fractionation for several reasons: (1) None of the standard materials exhibited a variation $>\pm 1.5\%$ about their mean $\delta^{13}\text{C}$ values, although weighted averages of the $\delta^{13}\text{C}$ obtained in a stepwise fashion are comparable with that derived from a single step combustion. (2) The spread of Allende isotopic measurements is 15 – 20% irrespective of grain size and is independent of whether any silicate has been removed from the sample. The release pattern of the

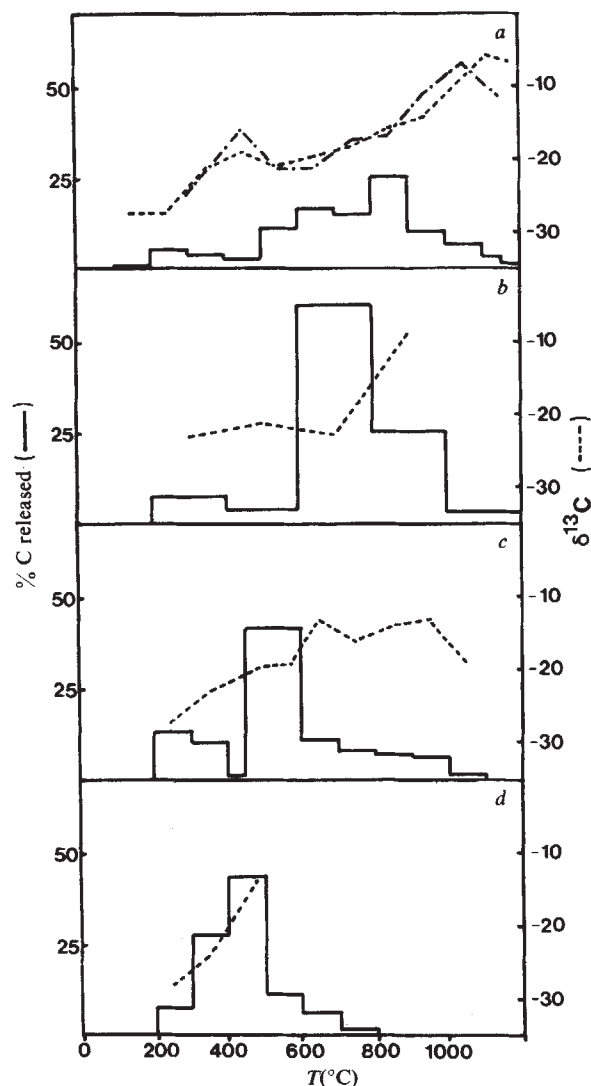


Fig. 2 Per cent carbon release and $\delta^{13}\text{C}$ composition from various size fractions and treatments of Allende. *a*, $>50\ \mu\text{m}$ bulk duplicate analysis, $\Sigma\delta^{13}\text{C} = -17.2, -17.1\%$, $\Sigma\text{C} = 0.21\%$, only one carbon release profile is shown; *b*, $<50\ \mu\text{m}$ bulk, $\Sigma\delta^{13}\text{C} = -18.8\%$, $\Sigma\text{C} = 0.27\%$; *c*, $<25\ \mu\text{m}$ bulk, $\Sigma\delta^{13}\text{C} = -18.9\%$, $\Sigma\text{C} = 0.37\%$; *d*, partially demineralized by HF/HCl, $\Sigma\delta^{13}\text{C} = -17.4\%$, $\Sigma\text{C} = 1.89\%$.

carbon from the sample is, however, governed by the physical condition of the sample, being easier to combust in the absence of minerals and much more difficult to release from the $>50\ \mu\text{m}$ fraction. The opportunity for kinetic isotope effects in the latter sample, which was analysed in duplicate, is obviously greater but clearly not excessively manifest. (3) None of the undemineralized meteorite samples show a regular and continuous increase in $\delta^{13}\text{C}$ with temperature. In the samples of $>50\ \mu\text{m}$ and $<50\ \mu\text{m}$, troughs occur in the isotope release pattern over the $500\text{--}700^\circ\text{C}$ interval. The $<25\ \mu\text{m}$ sample exhibited two distinct plateaus corresponding to peaks in the carbon release. (4) Stepwise combustion applied to other meteorite samples gives rise to complex release profiles which cannot be explained by simple kinetic fractionation¹⁰.

We suggest that for Allende, different forms of carbon are being oxidized or decomposed as temperature increases. Undoubtedly, some kinetic isotopic fractionation occurs, but this is probably within a few per mil, based on evidence from our stepwise combustion of individual carbonaceous substances. The profile obtained for the $<25\ \mu\text{m}$ fraction provides the best estimate of the isotopic compositions of the carbon phases (Fig. 2). We have estimated the isotopic compositions on two criteria:

first, the carbon shows sharp release profiles and second for the most part these profiles appear to correspond with plateaus in the carbon isotope ratio. The initial low temperature release (below 400°C) shows a rapid increase in isotopic ratio to the first plateau, perhaps as a result of mixing of two or more phases having only slightly different oxidation temperatures. The mid-temperature oxidizable phase, which contributes the bulk of the carbon is released over the interval $450\text{--}600^\circ\text{C}$ and has a mean $\delta^{13}\text{C}$ of $\sim -20.3\%$. The high temperature carbon in the $<25\ \mu\text{m}$ fraction is liberated above 600°C and has the heaviest carbon encountered in this size fraction and averages out at -13.4% . Higher $\delta^{13}\text{C}$ values may be encountered in coarser size fractions (Fig. 2) although these could reflect an enhanced kinetic isotope effect.

In the duplicate $>50\ \mu\text{m}$ sized fractions of Allende, a reduction in the carbon isotopic ratio between 500 and 600°C was observed. Such a pattern of isotopic release could be explained as the existence of a further, possibly low abundance, carbonaceous phase possessing a very negative isotopic composition. During stepwise combustion of $<4\ \mu\text{m}$ Allende matrix samples¹¹, the $\delta^{15}\text{N}$ plummets to -53% for the $1.28\ \text{p.p.m.}$ of nitrogen liberated during the 490°C step and correlates strongly with the release of a heavy xenon component suggesting the addition of an isotope of nucleosynthetic origin¹¹. The admix of nearly pure ^{14}N of presolar origin into Solar System materials has previously been discussed by several workers¹²⁻¹⁴. There may be no *a priori* significance between compatible release of nuclear components and a possible low $\delta^{13}\text{C}$, but the coincidence merits further study.

The carbon release and isotope patterns of the various fractions studied allow us to draw some conclusions regarding the nature and location of the phases believed to exist. First, we note that the mid-temperature phase ($<25\ \mu\text{m}$ nomenclature; Fig. 2c) which constitutes most of the carbon in the meteorite is of variable temperature release when other fractions are considered. Coarse sieve fractions do not release the phase completely until 900°C , but our partially demineralized sample and the AMII fraction of Ott *et al.*² suggest that in its free form this material combusts between 300 and 600°C . Such a reaction temperature is above the value found for a type I terrestrial kerogen, but well below that observed for NBS graphite and lends credence to the interpretation that the major carbon phase in Allende is a highly cross-linked, possibly aromatic, polymer, although some allotrope of carbon such as carbyne or amorphous carbon cannot be ruled out. We thus concur with the interpretation of Simmonds *et al.*¹⁵ that carbon in Allende is in a macromolecular form unlike graphite. We are unable to comment on the oxidation properties of "the interwoven ribbon-shaped packets of graphite layer plains" of Smith and Buseck¹⁶ or poorly crystalline graphite¹⁷. Although direct comparisons are difficult because of sample differences, the mid-temperature phase is most probably that discussed extensively by Frick and Pepin in relation to noble gases and nitrogen⁹. The mid-temperature carbon phases must be associated with a portion of the meteorite which can be easily crushed and available for oxidation; the evidence seems to point strongly towards the fine-grained olivine matrix which has been shown to have carbonaceous surface deposits^{17,18}. Second, the high temperature carbon phase of the $<25\ \mu\text{m}$ fraction does not show a pronounced temperature variation when other sieved undemineralized fractions are considered. Thus the heaviest carbon in the sample probably resides within the silicate minerals and corresponds to the so-called 'protected phase' of Frick and Pepin⁹. Once the silicate minerals are partially or completely removed, the carbon becomes oxidizable below 700°C suggesting that it is neither a highly crystalline nor ordered graphite. Third, the very low temperature phase released below 400°C in all undemineralized samples is clearly available to oxidation. Although precautions were taken, it cannot be ruled out that a proportion of terrestrial contamination is involved. Finally, we perceive, as previously reported by Chang *et al.*⁶ an increase in carbon content and a slight shift

towards lighter carbon as particle size decreases. These trends may be accounted for by an additional availability of the low and medium temperature phases and concurrent with their occurrence in fine-grained, easily crushed matrix.

We conclude that stepwise combustion of bulk and demineralized samples provides an easy and relatively contamination-free method of identifying the carbon isotopic composition of phases within Allende. The use of the technique in conjunction with noble gas and other volatile element analyses and its application to model materials should prove highly informative.

We thank the SERC for support, Professor E. Anders for stimulating our interest in stepwise combustion, the members of the Planetary Sciences Unit, Cambridge for helpful discussions and Mr I. Gilmour for the kerogen sample.

Received 17 February; accepted 22 April 1982.

1. Anders, E. *Proc. R. Soc.*, **A374**, 207–238 (1981).
2. Ott, U., Mack, R. & Chang, S. *Geochim. cosmochim. Acta* **45**, 1751–1788 (1981).
3. Frick, U. & Pepin, R. O. *Earth planet. Sci. Lett.* **56**, 45–63 (1981).
4. Clayton, R. N. *Science* **140**, 192–193 (1963).
5. Grady, M. M., Swart, P. K. & Pillinger, C. T. *Lunar planet. Sci.* **13**, 277–278 (1982).
6. Chang, S., Mack, R. & Lennon, K. *Lunar planet. Sci.* **9**, 157–159 (1978).
7. Mühle, K., Schnikow, B. L., Stiehl, G. & Winkler, I. *Chem. Erde* **39**, 27–36 (1980).
8. Grady, M. M., Swart, P. K. & Pillinger, C. T. *Meteoritics* **16**, 319 (1981).
9. Frick, U. & Pepin, R. O. *Lunar planet. Sci.* **11**, 303–305 (1980).
10. Swart, P. K., Grady, M. M. & Pillinger, C. T. *Lunar planet. Sci.* **13**, 788–789 (1982).
11. Frick, U. & Pepin, R. O. *Earth planet. Sci. Lett.* **56**, 64–81 (1981).
12. Kung, C. C. & Clayton, R. N. *Earth planet. Sci. Lett.* **38**, 421–435 (1978).
13. Thiemans, M. H. & Clayton, R. N. *Earth planet. Sci. Lett.* **55**, 363–369 (1981).
14. Geiss, J. & Bochsler, P. *Geochim. cosmochim. Acta* **46**, 529–548 (1982).
15. Simmonds, P. G., Bauman, A. J., Bollin, E. M., Gelpi, E. & Oró, J. *Proc. natn. Acad. Sci. U.S.A.* **64**, 1027–1034 (1967).
16. Smith, P. P. K. & Buscek, P. R. *Science* **212**, 322–324 (1981).
17. Bauman, A. J., Devaney, J. R. & Bollin, E. M. *Nature* **241**, 264–267 (1973).
18. Green, H. W., Radcliffe, S. V. & Heuer, A. H. *Science* **172**, 936–939 (1971).

Rainfall acidity in northern Britain

D. Fowler*, J. N. Cape*, I. D. Leith*, I. S. Paterson†, J. W. Kinnaid† & I. A. Nicholson†

* Institute of Terrestrial Ecology, Bush Estate, Penicuik, Midlothian EH26 0QB, UK

† Institute of Terrestrial Ecology, Hill of Brathens, Glassel, Banchory, Kincardineshire AB3 4BY, UK

The acidity of rain has attracted considerable international interest during the past decade. An area of enhanced rainfall acidity with its focus in the Low Countries of north-west Europe was identified by the European air chemistry study¹ for the period 1955–68, and widespread concern arose as a result of observations that fish populations were decreasing in areas of Scandinavia subject to acid rain². We report here measurements of acidity and related anion concentrations in rain collected at 16 sites in northern Britain during the period 1978–80. Estimates are provided of annual deposition of acidity in rain, the proportion of this acidity attributable to sulphuric and nitric acids, and an example of the influence of local concentrations of pollutant gases on rainfall acidity in cities.

The term 'acid rain' refers to rain more acid than water in equilibrium with atmospheric concentrations of carbon dioxide ($\sim$ pH 5.6, $2.5 \mu\text{Equiv. H}^+ \text{l}^{-1}$). Although several factors are involved, the areas in Scandinavia with significantly altered fish populations are subject to rainfall acidities of 30–60 $\mu\text{Equiv. H}^+ \text{l}^{-1}$ (pH 4.5–4.2). Fish mortality in acid water is caused mainly by problems with body salt regulation, the increased solubility of aluminium compounds in acid water being strongly implicated in these effects³. Evidence for links between acid precipitation and fish mortality is not restricted to Scandinavia, as similar effects have been reported for eastern North America⁴ and Scotland⁵.

In the United Kingdom there have been few detailed and systematic observations of rain chemistry in recent years, des-

pite early interest in the subject⁶. Gorham, in 1958, showed that areas of north-west England received appreciable inputs of pollutants, including acidity, from industrial areas to the south-east⁷. Ten years later, Stevenson reported concentrations of all the major ions in rain except hydrogen for eight sites in Britain⁸. In 1978 and 1979, Martin and Barber showed that rain in central and eastern England was 'acidic', with annual average hydrogen ion concentrations of 40–80 $\mu\text{Equiv. H}^+ \text{l}^{-1}$ (pH 4.4–4.1)^{9,10}.

At each of the 16 sites (Fig. 1) rain was collected using a 20 cm diameter Pyrex glass funnel mounted 2 m above ground, each funnel surrounded by a ring of spikes to discourage birds from perching on the rim. Pyrex glass has been shown to influence concentrations of sodium and potassium ions through leaching, and may underestimate acidity¹¹ but the surface properties are likely to remain more consistent over prolonged periods in the field than plastic funnels which deteriorate with time. Rain water flowed into darkened polypropylene bottles through a small piece of nylon gauze inserted into the neck of the funnel to prevent the collection of insects and larger debris. Although preservatives were not added, there was only very occasional evidence of algal growth occurring between the monthly collections of samples, which were sent by post to the laboratory where they were stored at 4 °C before being analysed. There remains the possibility of chemical changes occurring in the sample through biological activity during the intervals between collection. However, Galloway and Likens¹¹ showed that there were no significant changes in concentrations of the major ions, except chloride, over a 7 month storage period at 21 °C for a sample at pH 4. A continually exposed funnel rather than an automatic wet-only sampler was used as there were problems of reliability and contamination with the automatic samplers available in 1977. The contribution of dry deposition



Fig. 1 Weighted mean acidity of rain for the period 1978–80 in northern Britain from equation (1) (●, rural sites; ○, urban site).

Table 1 Site details and acidities, predicted* and measured, 1978–80

Site name	Map ref. (O.S.)	Altitude (m)	Measured (M) acidity ($\mu\text{Equiv. l}^{-1}$)	Predicted (P) acidity ($\mu\text{Equiv. l}^{-1}$)	P–M ($\mu\text{Equiv. l}^{-1}$)
Rural sites					
Bettyhill	NC 708 624	30	25	21	–4
Inverpolly	NC 075 145	30	21	19	–2
Nigg	NH 795 738	5	27	34	+7
Broadford	NG 628 247	15	21	24	+3
Forest of Deer	NJ 976 510	100	60	61	+1
Achnagoichan	NH 914 082	300	37	42	+5
Banchory	NO 680 988	140	59	57	–2
Lochnagar	NO 274 858	600	52	50	–2
Torlundy	NN 147 773	30	33	32	–1
Faskally	NN 919 600	150	50	44	–6
Lepinmore	NS 005 915	300	37	32	–5
Bush	NT 246 638	185	48	49	+1
Whiteadder†	NT 663 633	230	57	58	+1
Redesdale	NY 833 954	260	59	57	–2
Waterhead‡	NX 752 836	200	33	36	+3
Urban site					
Glasgow	NS 608 687	100	55	39	–16

* From equation (1).

† 1979–80 only.

‡ 1980 only.

to the ionic composition of rainfall collected in our gauges has been independently estimated, and is discussed later.

Measurements of pH and the concentrations of sulphate and magnesium ions were made on samples collected during 1978 and 1979; since 1980 the range of analyses has included concentrations of nitrate, chloride, sodium, potassium and ammonium ions (J.N.C. and D.F., work in preparation). pH was measured taking the necessary precautions for samples of

low ionic strength¹² using standard glass and calomel electrodes. Chloride concentrations were measured using an ion-selective electrode. Sulphate analysis was by the continuous-flow thoron method¹³ with aqueous isopropanol as solvent. Measurement of nitrate was by reduction to nitrite followed by complexation with sulphanilamide/*N*-1-naphthylethylenediamine dihydrochloride¹⁴ in a continuous-flow system, and ammonium was measured as ammonia using continuous flow through an ammonia-selective electrode. Magnesium and calcium concentrations were determined by atomic absorption spectroscopy (after addition of lanthanum) and sodium and potassium concentrations by flame emission spectroscopy.

Magnesium ion concentrations were used to estimate the sea-salt contribution to total ionic composition. Fifteen of the 16 rain-collecting sites were located in rural areas remote from local sources of air pollutants, with one site located in central Glasgow (Fig. 1). They provided an approximately uniform coverage of northern Britain (Table 1).

The surfaces of rain collectors are contaminated by dry deposition of gaseous SO₂ and NO₂ and particulate sulphate. As a result the deposition of sulphur and hydrogen ions in rain is overestimated, their estimates in southeastern Scotland being exaggerated by ~10% and in northern and western Scotland by 2–5%. The contamination was quantified from funnel washings following dry days and from laboratory experiments in which funnels were washed after exposure to SO₂ (unpublished data). Collectors sited at 2 m above ground, to limit contamination by surface-derived material, capture 5–10% less rain than identical collectors at ground level¹⁵. However, for estimates of annual deposition this error can be minimized by combining the solute concentrations found in the network with the amounts of rain recorded by the Meteorological Office.

The weighted mean acidity of rain was least in the north-west of northern Britain (20 $\mu\text{Equiv. H}^+ \text{l}^{-1}$, pH 4.7) and greatest in the east and south of northern Britain (60 $\mu\text{Equiv. H}^+ \text{l}^{-1}$, pH 4.2) (Fig. 1). The regional gradient, which is relatively simple, is best described by the following quadratic relationship between National Grid coordinates and hydrogen ion concentrations:

$$[\text{H}^+] (\mu\text{Equiv. l}^{-1}) = 3.80 \times 10^{-6} E^2 + 7.05 \times 10^{-7} E.N. - 3.16 \times 10^{-6} N^2 - 1.06 \times 10^{-2} E + 4.32 \times 10^{-2} N - 118 \quad (1)$$

where *N* and *E* are four-figure northings and eastings from Ordnance Survey maps.

This empirical equation accounts for 94% of the variation,



Fig. 2 Annual deposition of hydrogen ions ($\text{kg H}^+ \text{ha}^{-1}$) in rain over northern Britain.

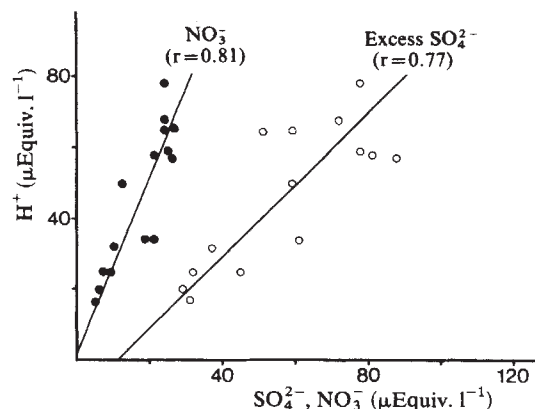


Fig. 3 Relationship between hydrogen ion concentrations in rain collected at 15 rural sites in northern Britain for the period 1978–80, and concentrations of nitrate (●) and non-marine (excess) sulphate (○) ions.

the average uncertainty of the predicted values for the locations marked in Fig. 1 being $\pm 9\%$ (Table 1).

As the regional pattern has been established from collections made in rural locations, it is of interest to consider local variations associated with major industrial and residential areas. To provide an estimate of the 'city influence' the Glasgow site at which the annual average SO_2 concentration is $\sim 80 \mu\text{g m}^{-3}$ (ref. 16) was included in the study. The regional pattern predicts a rainfall acidity in Glasgow equivalent to $39 \mu\text{Equiv. H}^+ \text{l}^{-1}$ (pH 4.45) whereas an acidity of $55 \mu\text{Equiv. H}^+ \text{l}^{-1}$ (pH 4.26) was actually recorded, indicating that $\sim 30\%$ of the acidity at that site was of local origin. The difference between the predicted regional value and measured local value exceeds the largest difference between predicted and measured values for the rural network sites by a factor of 2 (Table 1). The excess sulphate concentration (that sulphate remaining when sea-derived sulphate has been subtracted from the total) is also large in Glasgow, being approximately twice the regional pattern estimate (92 compared with $43 \mu\text{Equiv. l}^{-1}$). This suggests that some of the locally derived sulphate does not contribute to acidity, but seems to be associated with ammonium ions, for which the Glasgow concentration was over twice the average for northern Britain (43 compared with $19 \mu\text{Equiv. l}^{-1}$).

When the hydrogen ion concentrations for rain falling at the centre of each 10 km square of the National Grid are combined with the Meteorological Office's 30 yr data for amounts of precipitation, it can be seen that the deposition of hydrogen ions is strongly influenced by height above sea level (Fig. 2), an effect mainly attributable to the relation between amounts of rain and altitude over northern Britain. Rainfall during the period of study exceeded the 30-yr average by $\sim 10\%$ ¹⁷. Rainfall chemistry is not independent of altitude, fine rain often encountered at high altitudes having characteristically larger ionic concentrations¹⁸. Single rain events show a decrease in concentration with the amount of rain^{19–21}. As the combined effect of these two processes is uncertain, average rainfall chemistry has been assumed to be constant with altitude. It is clear from Fig. 2 that rainfall in the west more than compensates for smaller average acidities, so that although most of the region has a hydrogen ion input in rain of $\sim 0.5 \text{ kg H}^+ \text{ha}^{-1} \text{yr}^{-1}$, there are areas with much larger inputs where annual hydrogen ion deposition may exceed $1 \text{ kg ha}^{-1} \text{yr}^{-1}$. For a consideration of effects the degree of acidity is important, but the quantity of acid deposited may be equally important. Henriksen²² suggested that the effects of the direct inputs of acidity in rain, when acidifying bodies of freshwater, could be likened to a large-scale titration. The deposition pattern shown in Fig. 2 identifies areas of northern Britain where inputs of acidity are large and comparable with areas of Norway where fish populations have decreased markedly³.

The dominant anions in rain are sulphate, nitrate and chloride. Seawater is a significant source of chloride and sul-

phate ions in rain, but makes a negligible contribution to nitrate or hydrogen ion concentrations. Our evidence suggests that most chloride has a marine origin, and therefore does not contribute to acidity, although non-marine chloride has been detected at a site in central Scotland close to known sources of gaseous HCl (unpublished data). The proportion of marine sulphate varies from $>60\%$ at coastal western sites to 10% in the east. The two major anions associated with acidity are therefore nitrate and excess sulphate, each of which is strongly correlated with hydrogen ion (Fig. 3). Similar correlations have been noted by others^{23,24}. There are two extreme interpretations of these correlations. On the one hand, all of the nitrate could be associated with hydrogen ions, in which case 38% of the acidity would be nitric acid, the remaining 62% being sulphuric acid. At the other extreme there is sufficient excess sulphate to account for all of the measured acidity as sulphuric acid. However, because ammonium sulphate is frequently the major component of sub-micrometre aerosols in the United Kingdom²⁵ and because sulphate concentrations in rain are closely associated with aerosol sulphate concentrations²⁶, a significant proportion of the excess sulphate in rain is likely to occur as ammonium sulphate. This suggests that at least some of the acidity is associated with nitrate ions and that the second interpretation is unlikely to be tenable. On average an increase in hydrogen ion concentration of $10 \mu\text{Equiv. l}^{-1}$ was associated with increases in nitrate and sulphate concentrations of 4 and $10 \mu\text{Equiv. l}^{-1}$ respectively, which suggests an average contribution to acidity of 29% by nitric acid and 71% by sulphuric acid. This is consistent with conclusions reached for the USA²⁷ and Scandinavia²⁸.

The acidity of rain falling in the east and south-east of northern Britain is comparable with that in areas of Scandinavia and North America where populations of fish have been severely depleted^{3,4}. The regional pattern of acidity shows no correlation with population density nor with the distribution of major sources of sulphur dioxide and nitrogen oxides within the region²⁹. Associated studies have shown the acidity to be highly episodic with a complex frequency distribution, and that the most acid episodes (pH < 4.0) arise from air masses that have had a trajectory over a major industrial region of the UK or Europe (D.F. and J.N.C., work in preparation). This implies that the scales of transport of the substances which acidify rain are much larger than the scale of northern Britain, and that sources within this region only influence the regional distribution of acidity in rain to a minor extent.

We thank the Forestry Commission, Nature Conservancy Council, Department of Agriculture and Fisheries for Scotland, Hill Farming Research Organisation, Freshwater Biological Association, Glasgow University, Ministry of Agriculture, Fisheries and Food, and Glasgow District Council for assistance with the collection of rain samples, Mr J. N. R. Jeffers and Professor F. T. Last for their helpful comments, and Miss M. McPherson for technical assistance. This work was financed by the NERC, the EEC and the UK Department of the Environment.

Received 16 December 1981; accepted 26 March 1982.

- Sweden's case to the U.N. Conf. on the Human Environment (Royal Ministry of Foreign Affairs, Stockholm, 1971).
- Leivestad, H., Hendrey, G., Muniz, I. P. & Snekvik, E. in *SNSF Report FR6/76* (ed. Braekke, F. H.) 87–111 (SNSF, Oslo 1976).
- Muniz, I. P. & Leivestad, H. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 84–92 (SNSF, Oslo, 1980).
- Harvey, H. H. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 93–98 (SNSF, Oslo, 1980).
- Harriman, R. & Morrison, B. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 312–313 (SNSF, Oslo, 1980).
- Lawes, J. B., Gilbert, J. H. & Pugh, E. *Phil. Trans. R. Soc. II*, 431–577 (1861).
- Gorham, E. *Phil. Trans. R. Soc. B241*, 147–178 (1958).
- Stevenson, C. M. Q. *J.R. met. Soc.* **96**, 56–70 (1968).
- Martin, A. & Barber, F. R. *Atmos. Envir.* **12**, 1481–1487 (1978).
- Martin, A. *Sci. Total Envir.* **13**, 119–130 (1979).
- Galloway, J. N. & Likens, G. E. in *Proc. 1st Int. Symp. Acid Precipitation and Forest Ecosystem* (eds Dochinger, L. S. & Seliga, T. A.) 137–155 (USDA Forest Service General Technical Report NE-23, 1976).
- Galloway, J. N., Crosby, B. J. & Likens, G. E. *Limnol. Oceanogr.* **24**, 1161–1165 (1979).
- Persson, G. A. *Air Water Pollut. Int. J.* **10**, 845–852 (1966).

14. Eggleton, A. E. J. & Atkins, D. H. *Results of the Tees-side Investigation*, 165 (AERE-R 6983, HMSO, London 1972).
15. Robinson, A. C. & Rodda, J. C. *Met. Mag.* **98**, 113–120 (1969).
16. *National Survey Smoke and Sulphur Dioxide* (Warren Spring Laboratory, Department of Industry, 1979).
17. *Isopercental Rainfall Maps for U.K. 1978–1980* (UK Meteorological Office, to be published).
18. Tomlinson, G. H., Brouzes, R. J. P., McLean, R. A. N. & Kadlecsek, J. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 134–137 (SNSF, Oslo, 1980).
19. Granat, L. *Atmos. Envir.* **12**, 413–424 (1978).
20. Hicks, B. B. & Shannon, J. D. *J. Am. met. Soc.* **18**, 1415–1420 (1979).
21. Fowler, D. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 22–32 (SNSF, Oslo, 1980).
22. Henriksen, A. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 68–74 (SNSF, Oslo, 1980).
23. Marsh, A. R. W. *Atmos. Envir.* **12**, 401–406 (1978).
24. Glover, G. M., Kallend, A. S., Marsh, A. R. W. & Webb, A. H. in *Effects of Acid Precipitation on Terrestrial Ecosystems* (eds Hutchinson, T. C. & Havas, M.) 95–110 (Plenum, New York, 1980).
25. Heard, M. J. & Wiffen, R. D. *Atmos. Envir.* **3**, 337–340 (1969).
26. Garland, J. A. *Atmos. Envir.* **12**, 349–362 (1978).
27. Galloway, J. N. & Likens, G. E. *Atmos. Envir.* **15**, 1081–1085 (1981).
28. Overrein, L. N., Seip, H. M. & Tollan, A. (eds) *Final Rep. SNSF Project, 1972–1980*, 23–32 (SNSF, Oslo, 1980).
29. Fowler, D. & Cape, J. N. in *Effects of Gaseous Air Pollutants in Agriculture and Horticulture* (eds Unsworth, M. H. & Ormrod, D. P.) (Butterworths, London, in the press).

Geomagnetic secular variation as a precursor of climatic change

V. Courtillot, J. L. Le Mouél & J. Ducruix

Département des Sciences de la Terre, Université Paris VII et Institut de Physique du Globe, Université Paris VI, 2 Place Jussieu, 75230 Paris Cedex 05, France

A. Cazenave

Groupe de Recherches de Géodésie Spatiale, CNES, Toulouse, France

Long period trends in climate are usually associated with solar disturbance¹. For example, a close similarity has been demonstrated^{2,3} between variations in the length of day (LOD) with periods greater than about 10 yr and trends in several climatic indices over the past 150 yr. We point out here a correlation between variations in the Earth's magnetic field, the Earth's rotation rate and some climatic indicators, thus suggesting a possible long term influence of core motions on climate. We suggest that geomagnetic secular variation can be used to forecast a climatic change in the 1990s.

A good correlation has been established^{2,3} between LOD variations and either global temperature θ (correlation coefficient 0.85; θ lagging LOD variations by ~ 5 yr, see Fig. 1) or the long period atmospheric and oceanic excitation function ψ^T related to near surface winds, changes in the atmospheric mass distribution and eustatic changes in sea level (correlation coefficient 0.75; ψ^T lagging LOD variations by ~ 15 yr see Fig. 9.6 in ref. 4; ψ^T is computed from ground level

pressure sea-level data under the geostrophic assumption). Periods of increasing zonal circulation and increasing global surface temperature are found to correspond to periods of Earth acceleration, while periods of decreasing zonal circulation and decreasing global surface temperature correspond to periods of Earth deceleration. It has been further observed³ that the excitation function ψ^T represents only about 10% of the amount necessary to explain the LOD variations. This led to the conclusion that it is not possible as yet to establish causal relationships between these phenomena and that one cannot decide whether (1) atmospheric circulation causes the long period changes in LOD, or (2) the fluctuations are both a consequence of a third phenomenon, or (3) the fluctuations in LOD cause the observed variations in the atmospheric circulation.

Although A.C. and Lambeck³ note the problems related to phase lag between θ (or ψ^T) and LOD and the insufficient amplitude of the excitation function, they tend to reject hypothesis (3) because "the total mechanism for exciting the LOD changes remains to be explained".

More recently, Lambeck⁴ has reviewed the long period or decade fluctuations in the Earth's rotation and discussed their geophysical origin. He notes that the role of the core is central, for it is the only sufficiently mobile part of the Earth with sufficient mass to modify the rotation by the observed amounts on that time scale. Out of the mechanisms proposed for core-mantle coupling, pressure or inertial coupling due to the ellipticity of the core boundary, topographic coupling due to bumps on the core-mantle interface and viscous coupling all seem to be inadequate^{4,5}. Only electromagnetic core-mantle coupling, as suggested by Bullard *et al.*⁶, survives the screening process despite severe uncertainties regarding the conductivity of the lower mantle and the core motions. The key observations required to resolve these uncertainties come from observations of the secular variation, in particular the rate of westward drift, of the geomagnetic field. The magnetic observations available to Lambeck did suggest some kind of correlation between LOD and geomagnetic variations, but the evidence was inconclusive.

The situation has been improved with the observation⁷ that the secular variation of the geomagnetic field had undergone sharp accelerations of global extent. The last and best studied such phenomenon which occurred around 1970 is described in refs 8, 9, 14. The secular acceleration can first be used to infer an upper bound on the electrical conductivity of the lower mantle which is found to be of the order of a few hundred $\Omega^{-1} \text{ m}^{-1}$ (refs 10, 11). This parameter is essential in any computation related to electromagnetic core-mantle coupling.

The correlation⁷ between secular variation accelerations and extrema in LOD fluctuations has recently been put on a firmer and more systematic basis¹² with the demonstration of a good correlation between LOD variations and the secular variation of declination, observed for instance in European observatories where the longest records are available: the correlation for the period 1865–1975 is far more convincing than any previously reported (Fig. 1). The correlation coefficient is found to be about 0.8, with the magnetic variations leading LOD fluctuations by about 10 yr. J.L. Le M. and V.C.¹³ suggest that a model of accelerated westward drift of the upper core based on the original model of Bullard *et al.*⁶ accounts both qualitatively and quantitatively for the observations. In this model, the Earth's mantle, an upper core layer and the lower core are all electromagnetically coupled. A sudden torque applied at the time of a secular variation acceleration on a ~ 100 km thick upper core layer leads to a very fast response of westward drift as observed at the surface, and to a fast response of the lower core in the opposite sense. The lower core finally drags the mantle with a longer time-constant, of the order of 10 yr, as is typically observed in the decade fluctuations of LOD. Thus, these decade fluctuations are probably caused by rearrangements in core motions. Having found a mechanism which can excite LOD changes, it becomes natural to accept the hypothesis that fluctuations in LOD in turn are responsible for changes in atmospheric circulation (although the hypothesis of an indepen-

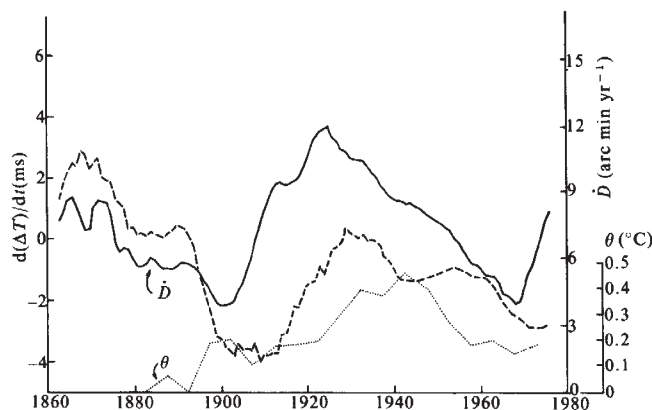


Fig. 1 Plot of the secular variation of geomagnetic declination Δ (solid curve), of the excess length of day $d(\Delta T)/dt$ (dashed curve) and of successive 5 yr averages of the temperature θ over the whole Earth (expressed as departures from the means for 1880–84, after Fig. 18.2 of ref. 1, dotted curve).

dant common cause cannot be rejected). The observed phase lag between ψ^T and LOD and the amplitude of ψ^T are then at least qualitatively accounted for. There is an attractive causal chain whereby, over the decade period range, angular momentum is transferred from the core to the mantle and finally to the atmosphere (although details of the transfer mechanism remain to be studied).

Lambeck and A.C.³ point out that LOD variations can be used as an indicator of future climatic trends. Lamb (ref. 1, p. 710) in his review of scientifically based climate forecasts notes that variations are commonly associated with long term variations of solar disturbance. He notes a considerable measure of agreement between the various forecasts, with a trend towards colder climates with weakened general atmospheric circulation which is expected to continue into the twenty-first century, in some cases with a sharp further cooling about 1980

Received 9 February; accepted 21 April 1982.

1. Lamb, H. H. *Climate, Present, Past and Future*, Vol. 2 (Methuen, London, 1977).
2. Cazenave, A. thesis, Univ. Toulouse (1975).
3. Lambeck, K. & Cazenave, A. *Geophys. J. R. astr. Soc.* **46**, 555–573 (1976).
4. Lambeck, *The Earth's Variable Rotation. Geophysical Causes and Consequences* (Cambridge University Press, London, 1980).
5. Rochester, M. G. *J. AGU Bull.* **43**, 115 (1979).
6. Bullard, E. C., Freedman, C., Gellman, H. & Nixon, J. *Phil. Trans. R. Soc. A* **243**, 67–92 (1950).
7. Courtillot, V., Ducruix, J. & Le Mouél, J. L. *C.R. heb. Séanc. Acad. Sci., Paris D* **287**, 1095–1098 (1978); **B **289**, 173–175 (1979).**

and easier conditions for a time in the first half of the next century. The geomagnetic results referred to here suggest that changes in secular variation can be used as indicators of future LOD fluctuations^{12,13} and, with an even longer time lead of 15–25 yr, as indicators of climatic changes, in particular changes in global surface temperature. The clearly established 1970 secular acceleration, which has now been maintained for a decade, suggests a correlated positive acceleration in LOD in the immediate future and the beginning of a period of steady increase in average global temperature around 1990 (± 5 yr). K. Brian (personal communication) has recently pointed out a possible correlation between the frequency of geomagnetic reversals and deep ocean temperature over the past 55 Myr, thus suggesting another (very) long term relationship between climate and the Earth's magnetic field.

This work is IPG contribution no. 582.

8. Ducruix, J., Courtillot, V. & Le Mouél, J. L. *Geophys. J. Roy. astr. Soc.* **61**, 73–94 (1980).
9. Ha Duyen, C., Ducruix, J. & Le Mouél, J. L. *C.R. heb. Séanc. Acad. Sci., Paris B* **293**, 157–160 (1980).
10. Achache, J., Courtillot, V., Ducruix, J. & Le Mouél, J. L. *Phys. Earth planet. Inter.* **23**, 72–75 (1980).
11. Achache, J., Le Mouél, J. L. & Courtillot, V. *Geophys. J. R. astr. Soc.* **65**, 579–601 (1981).
12. Le Mouél, J. L., Madden, T. R., Ducruix, J. & Courtillot, V. *Nature* **290**, 763–765 (1981).
13. Le Mouél, J. L. & Courtillot, V. *Phys. Earth planet. Inter.* **24**, 236–241 (1981); *J. geophys. Res.* (in the press).
14. Malin, S. R. C. & Hodder, B. M. *Nature* **296**, 726–728 (1982).

Spore-pollen evidence for early Oligocene high-latitude cool climatic episode in northern Canada

Geoffrey Norris

Department of Geology, University of Toronto, Toronto, Ontario, Canada M5S 1A1

Climatic cooling in the Eocene leading to markedly lower temperatures in the Oligocene has been documented in oxygen isotope studies in the North Pacific^{1,2}, New Zealand³, South Pacific⁴ and North Sea⁵. These lower temperatures in the Oligocene correspond to palaeoclimatic interpretations of North American leaf floras^{6,7} which suggest a profound cooling occurring towards the end of the Eocene, indicated also in the southeastern United States from studies on mid-Tertiary spores and pollen⁸. Recently, Collinson *et al.*⁹ have shown palaeobotanically that in southern England cooling occurred gradually starting in the latest early Eocene leading to two major periods of floristic change before the end of the Eocene. Thick Tertiary sections in the subsurface Mackenzie Delta region, Northwest Territories^{10,11} provide well-preserved Palaeogene palynofloras¹² discussed here which indicate a cooler climate in the Oligocene following warm-temperate conditions in the Eocene. This widespread early-middle Oligocene cool episode is thus represented in both high-latitude and lower-latitude floras and seems to have been of global extent, persisting until towards the end of the Oligocene, but its precise dating remains a problem. In northern Canada it is followed by an amelioration of the climate in the late Oligocene which persisted probably until the middle Miocene. The rate of cooling during the Eocene and Oligocene may have varied depending on locality, but once established the cooler Oligocene climate exerted a profound influence on biotic development.

The subsurface Mackenzie Delta region (lat. 69°N, long. 134°W) includes a thick Tertiary molasse pile^{11,13,14} deposited in the Richards Island Basin, representing several regressive depositional events formed by prograding deltas with marine deposits at the base. A relatively continuous Tertiary section is available here from the Palaeocene through to the Quaternary^{11,12,15,16} but a major unconformity removes much of the Palaeogene in the southern part of the basin. The Imperial Nuktak C-22 well in the northern part of the basin has a

relatively complete Tertiary section (Fig. 1) down to TD 12,650 ft in the middle Eocene part of the Richards Formation¹², a prodeltaic mudstone unit in excess of 5,000 ft thick which passes laterally southwards into the upper part of the Palaeocene-middle Eocene deltaic Reindeer Formation¹¹. The Richards Formation is largely middle to late Eocene on the basis of the *Haplophragmoides* foraminiferal assemblage¹¹ and the presence near the base¹² of *Pistillipollenites mcgregorii* Rouse and the dinoflagellate cysts *Glaphyrocysta ordinata* (Will.) and Dow) Stover and Evitt, *Cordosphaeridium gracile* (Eis.) Dav. and Will., and *Wetzelia* sp. cf. *W. hamptensis* Wilson. The precise placement of the Eocene-Oligocene boundary has not been achieved yet because ranges of terrestrial palynomorphs are uncertain at these high latitudes^{12,15,17}. It probably occurs almost 1,700 ft below the top of the Richards Formation (Fig. 1) if Rouse¹⁷ is correct in limiting the range of *Integricarpus* sp. (which probably does not occur below 9,500 ft in the Nuktak C-22 well) to the Oligocene. This interpretation is supported by the presence in the upper 1,700 ft of the Richards Formation of spore-pollen species from the putative Lower Oligocene of northeastern China¹⁸.

The terrestrial flora of the middle-late Eocene part of the Richards Formation is diverse and contains taxa which can be

LITHOSTRATIGRAPHIC UNIT	PALYNOLOGY ZONE	INFERRED AGE	DEPOSITIONAL ENVIRONMENT
NUKTAK FORMATION	LAEVIGATOSPORITES	PLIOCENE	NON-MARINE
	CHENOPODIPOLLIS	PLIOCENE OR LATE MIOCENE	RESTRICTED 500
	TSUGAEPOLLENITES	MIOCENE	MARINE SAND & MUD
	ERIPITES	LATE OIGOCENE	NON-MARINE 1800
	IVIK MB.	7810	MUD AND GRAVEL
RICHARDS FORMATION	OSMUNDACIDITES	EARLY OIGOCENE	NON-MARINE 3150
	INTEGRICARPUS	LATE EOCENE	COASTAL PLAIN 4518 AND DELTAIC SANDS
	PESAVIS	MIDDLE EOCENE	NON-MARINE DELTA FRONT MUD AND SAND PASSING DOWN INTO PRODELTAIC MUDS
T.D. 12650			BRACKISH 10400 PRODELTAIC MUDS
			MARINE 11700 PRODELTAIC MUDS

Fig. 1 Tertiary stratigraphy of Imperial Nuktak C-22 well. Palynostratigraphy, zonal names and age according to ref. 12, but Eocene-Oligocene boundary may occur lower (see text for discussion). Lithostratigraphy after Young and McNeil¹¹. Numbers give the depth in feet.

attributed to the following extant genera as discussed in detail elsewhere^{8,12,15,19,20}: *Pinus*, *Sequoia*, *Metasequoia*, *Ulmus*, *Tilia*, *Castanea*, *Quercus*, *Tsuga*, *Pterocarya*, *Picea*, *Alnus*, *Betula* and *Sphagnum*. The latter four genera range into cool temperate or boreal regions at the present, but the remaining nine taxa are clearly temperate or warm temperate. This mixed palynoflora may be due to long-distance pollen transport from cooler—perhaps mountainous—regions of the northern Cordillera into warm temperate lowlands prevailing in the Mackenzie Delta region, a possibility supported by the sedimentology¹³. Mixing due to cavings in the cuttings samples may also introduce non-indigenous species into the samples, indicated by the lack of *Picea* and *Betula* in core samples compared with cuttings samples of the Richards Formation. Thus the indigenous palynoflora of the Richards Formation indicates temperate or warm temperate climates in the middle and late Eocene, possibly persisting into the earliest Oligocene. A sharp decrease in diversity of palynofloras and extinction of several species occurs within the prodeltaic mudstones of the early Oligocene part of the Richards Formation between 9,600 and 8,900 ft and is not associated with a major facies change. The Ivik Member of the superadjacent Kugmallit Formation contains a less diverse palynoflora of notably different composition, particularly in the upper part of the member, compared with the Richards Formation. Miospores attributed to extant taxa in the Ivik floras include *Sphagnum*, *Picea*, *Pinus*, *Lycopodium*, *Osmunda*, *Tsuga*, *Alnus*, *Betula*, *Carpinus* and *Pterocarya*. However, it is significant that a variety of fungal spores and hyphae, bryophytes and pteridophytes, *Sequoia*, *Ulmus*, *Tilia*, *Quercus*, *Castanea*, and other thermophilic angiosperms which are common in the Eocene—earliest Oligocene Richards Formation disappear in the lower half of the early—middle Oligocene¹² Ivik Member (below 6,500 ft) but reappear higher in the late Oligocene^{11,12} Arnak Member of the Kugmallit Formation accompanied by ericaceous pollen amongst others. The lower Ivik palynofloral changes occur in an interval of rhythmic alterations of mudstone and sandstone in medium-scale coarsening-upward trends typical of progradational deltas. No major facies change occurs in this interval. The lack of thermophilic angiosperm taxa in part of the early and perhaps middle Oligocene indicates a significant cool interval, possibly cooler than the climate of the present Great Lakes region where these taxa terminate their northern ranges in the pollen rain²¹. Their resurgence in the Arnak Member is accompanied by an overall average doubling of terrestrial species numbers. This cannot be attributed entirely to a change to alluvial or deltaic conditions from the delta-front conditions of the Ivik Member because other similar progradational cycles in the Richards Island Basin¹¹ are not characterized by such a change in palynofloral diversity (unpublished data). The increased diversity, however, may be related to an amelioration of climate as warm temperate conditions were re-established.

The reasons for this early to middle Oligocene cool interval are not clear. An explanation in terms of a change in obliquity of the ecliptic^{6,7} is not considered dynamically acceptable²², nor is it in accord with other floral and faunal evidence of pronounced Palaeogene seasonality indicating significant annual periods of winter darkness at high latitude²³. However, a cooler interval during the Oligocene appears to be of global extent, affecting oceanic water temperatures in both hemispheres¹⁻⁴ affecting the marine biota^{22,24}, and affecting high and low latitude terrestrial floras for at least 5 Myr. In Central Europe, however, the coolest mid-Tertiary terrestrial floras are late Oligocene²⁵, in contrast to the early and middle Oligocene cooling indicated in many regions as outlined above. In central British Columbia, Piel²⁶ reported humid warm temperate to sub-tropical climates for the putative early Oligocene Australian Creek Formation of central British Columbia based on the presence of miospores attributed to *Metasequoia*, *Picea*, *Pinus*, *Podocarpus*, *Acer*, *Betula*, *Carya*, *Castanea*, *Engelhardtia*, *Fagus*, *Fraxinus*, *Juglans*, *Liquidambar*, *Nyssa*, *Tilia* and *Ulmus/Zelkova*. Rouse and Mathews²⁷, however, have interpreted this assemblage to

be late early Oligocene representing a mixed mesophytic warm temperate forest, similar to those of southeastern United States or the upper Yangtze River Valley of China. The age of these palynofloras is contentious but may be indicated by the report of associated titanotheres teeth (a brontotherine genus, possibly *Megacerops* or *Brontotherium*) said to be late early Oligocene²⁷. On the other hand, the presence here of species of *Parviprojectus*, *Pesavis*, *Ctenosporites* and *Striadiporites* points to a correlation with the putatively middle and late Eocene *Pesavis* and *Integricorpus* zones in the Richards Island Basin¹² (Fig. 1). Climatic gradient may have strongly affected the differential attenuation of species ranges at different latitudes in the late Palaeocene, thus allowing the extension of arctic Eocene palynomorph species into the early Oligocene at lower latitudes. Until this matter is resolved, however, the possibility exists that the *Integricorpus* zone may be entirely early Oligocene and, therefore, that the climatic cooling indicated by palynofloral data in the upper part of the *Integricorpus* zone and higher in the western Arctic may have occurred later than the Eocene—Oligocene boundary.

This dramatic temperature drop in the Oligocene leads to a paradox²⁸. Cooler global air and water temperatures would lead to decreased evaporation rates on a global scale, but, nevertheless, moisture continued to be transported to Antarctica where ice continued to accumulate and reached sea level at the end of the Oligocene. Ice accumulation in the Arctic Ocean did not commence until some time in the Pliocene²⁹ when the Miocene temperate terrestrial floras were replaced by boreal tundra floras¹², and at which time a major episode of molluscan extinction occurred globally, probably in response to a lethal decline in temperature³⁰.

The palaeoceanographical changes at the Eocene—Oligocene boundary have been related²⁴ to cooler air temperatures and floral changes on land. O'Keefe's hypothesis²² for early Oligocene cooling postulates the development of tectite ring system around the Earth which cast a shadow on the winter hemisphere. This would lead to a temperature drop of ~20 °C and remain stable for several million years. Synchronous global cooling could also result from a variation in the solar output²³. Progress in understanding the mechanisms will depend in part on increased precision in dating Palaeogene terrestrial deposits. The exact timing of terrestrial floral changes from region to region must be determined and compared with biotic and isotopic changes in the oceans. If the floral and faunal changes are indeed synchronous globally, a pervasive and almost instantaneous mechanism must be sought. If the terrestrial floral changes accompanying cooling prove to be diachronous, a mechanism dependent on local or regional changes in geography, oceanic circulation, or tectonics may be implicated.

This work was supported by an operating grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada, a contract from the Department of Supply and Services, Ottawa and a NSERC strategic energy grant for Basin Analysis at the University of Toronto. Samples and data were provided by the Institute of Sedimentary and Petroleum Geology, Calgary, Austin and Cumming Exploration Consultants, and Imperial Oil Ltd. Dr Glenn Rouse provided valuable information and discussion.

Received 14 January; accepted 5 April 1982.

1. Douglas, R. G. & Savin, S. M. *Init. Rep. DSDP Leg. 32*, 509–520 (1975).
2. Savin, R. J. A. *Rev. Earth planet. Sci.* **5**, 319–355 (1977).
3. Devereux, I. *New Z. J. Sci.* **10**, 988–1101 (1967).
4. Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP Leg. 29*, 743–755 (1975).
5. Buchardt, B. *Nature*, **275**, 121–123 (1978); *Dansk. Met. Inst. Climatol. Pap.* **4**, 243–248 (1978).
6. Wolfe, J. A. *Am. Scient.* **66**, 694–703 (1978).
7. Wolfe, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30**, 313–323 (1980).
8. Frederiksen, N. O. *J. Paleont.* **54**, 728–739 (1980).
9. Collinson, M. E., Fowler, K. & Boulter, M. C. *Nature* **291**, 315–317 (1981).
10. Young, F. G., Myhr, D. W. & Yorath, C. J. *Geol. Surv. Can. Pap.* **76**–11 1–65 (1976).
11. Young, F. G. & McNeil, D. W. *Geol. Surv. Can. Bull.* (in the press).
12. Norris, G. *Geol. Surv. Can. Bull.* **340** (in the press).
13. Dixon, J. *Geol. Surv. Can. Pap.* **80–81**, 1–11 (1981).
14. Bowerman, J. N. & Coffman, R. C. *Can. Soc. petrol. Geol. Mem.* **4**, 649–662 (1975).

15. Staplin, F. L. *Bull. Can. petrol. Geol.* **24**, 117–36 (1976).
16. Ioannides, N. S. & McIntyre, D. J. *Geol. Surv. Can. Pap.* **80-1A**, 197–208, (1980).
17. Rouse, G. E. *Am. Ass. Strat. Palynol., Contr. Ser.* **5A**, 48–65 (1977).
18. Sung, Z. C. *et al. Nanjing Inst. Geol. Paleont., Chinese Acad. Sci* 1–177 (Kexue Chubanshe, Peking, 1978).
19. Gruas, Cavagnetto, C. *Mem. Soc. Geol. Fr.* **131**, 1–64 (1978).
20. Elisk, W. C. *Palaeontographica, Abt. B.* **149**, 90–111 (1974).
21. Webb, I. & McAndrews, J. H. *Mem. geol. Soc. Am.* **145**, 267–299 (1976).
22. O'Keefe, J. A. *Nature* **285**, 309–311 (1980).
23. McKenna, M. C. *Palaeogeogr., Palaeoclimatol. Palaeoecol.* **30**, 349–362 (1980).
24. Keigwin, L. D. *Nature* **287**, 722–725 (1980).
25. Hochuli, P. *Ann. Geol. Pays Hellen., Tome hors ser. fasc. II*, 515–523 (1979); *Beitr. Palaeont. Ostr.* **4**, 1–132 (1978).
26. Piel, K. M. *Can. J. Bot.* **49**, 1885–1920 (1971).
27. Rouse, G. E. & Mathews, W. H. *Bull. Can. petrol. Geol.* **27**, 418–445 (1979).
28. Frakes, L. A. *Climates Throughout Geologic Time*, 1–310 (Elsevier, Amsterdam, 1979).
29. Clarke, D. L. *Bull. geol. Soc. Am.* **82**, 3313–3324 (1971).
30. Stanley, S. M. & Campbell, L. D. *Nature* **293**, 457–459 (1981).

Carbon cycle changes of the Zechstein Sea: isotopic transition zone in the Marl Slate

M. Magaritz

Isotope Department, The Weizmann Institute of Science, Rehovot, Israel

P. Turner

Department of Geological Sciences, The University of Aston, Gosta Green, Birmingham B4 7ET, UK

The carbon isotope composition of carbonate rocks is one parameter used to study the carbon, sulphur and oxygen cycles through geological time^{1,2}. The understanding of these interrelated cycles is important in quantitative flux modelling of sedimentary rocks. The carbon isotope composition in the large reservoir of dissolved carbonate in ocean water is a result of mass balance between the amount of carbon in the reduced organic reservoir and the oxidized carbonate reservoir at any given stage. The reduced carbon reservoir preferentially sequesters ¹²C owing to biological fractionation processes. Any change in the removal ratio of oxidized/reduced carbon will be reflected by the change of the $\delta^{13}\text{C}$ value of the oceanic carbonate system. Here we present a detailed study of the change in ¹³C which took place at the base of the Upper Permian Zechstein sequence.

On the basis of measurements of carbon isotopes on rocks of various ages, early models assumed that the mean $\delta^{13}\text{C}$ (carbonate) values were constant^{3–6}. In recent years, however, descriptions of secular variation of $\delta^{13}\text{C}$ values in carbonate rocks during the Mesozoic^{7,8} and Cenozoic^{9,10} have become established. These secular variations are explained by large scale burial of organic material and are associated with ocean wide changes such as formation of anoxic conditions and sea-level fluctuation¹¹.

The same mechanism is used in modelling mass balance chemical and isotopic variation in the C–S–O system between geological periods^{1–3}. These mass balance models^{2,3} assume that high $\delta^{13}\text{C}$ values prevail for the whole of the Permian period¹² but, in fact, the detailed evidence indicates that only certain parts of the Permian are characterized by these high values¹³. The ¹³C maximum in carbonate rocks discussed here starts at the base of the Upper Permian Zechstein sequence and lasts up to the top of the third Zechstein cycle¹⁴ when fully marine conditions terminated. Because such a change also means changes in the amount of oxygen tied to the carbon system then it may be expected to be reflected in the oxygen cycle and the cycles of other elements to which oxygen is related (sulphur and iron).

Some of the carbonate rocks of the Permian period exhibit anomalous enrichment in ¹³C relative to other Phanerozoic

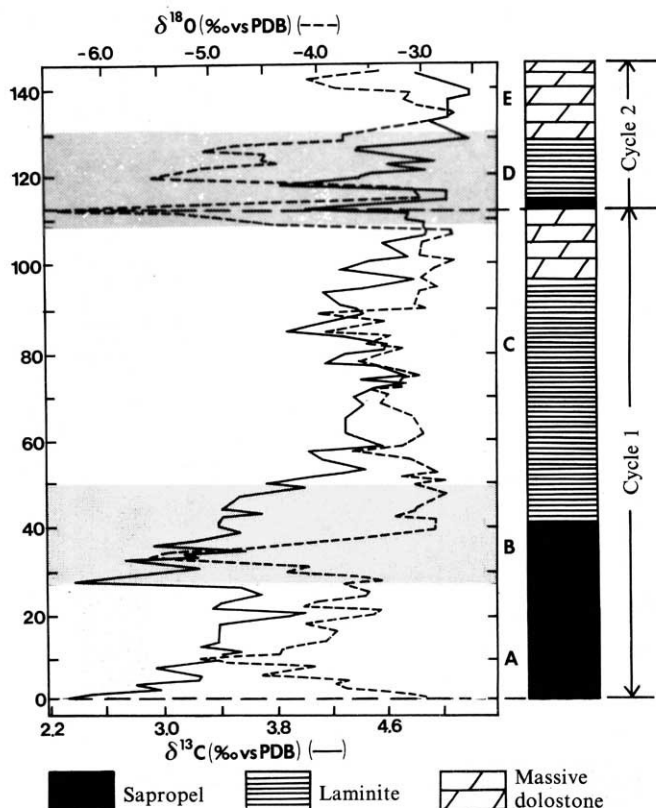


Fig. 1 Carbon and oxygen isotope composition of VT8 core dolomites. Note the spikes of low values of $\delta^{18}\text{O}$ which occur twice in zone A and zone B representing freshwater reflux of the Zechstein basin. $\delta^{13}\text{C}$ increases rapidly in zone B and remain constantly high in the upper part of the Marl Slate (zones C–E).

rocks. These high $\delta^{13}\text{C}$ values are observed even if one mixes all the Permian rocks analysed and compared them with rocks from other periods^{12,15}. The fact that some of the Permian carbonate rocks from various parts of the world reach values of up to +7‰ (relative to the normal fluctuation range –1 to +3‰)^{16–19} indicates that extreme changes in the C–O–S cycle took place. These changes had also been predicted from the anomalously low values of $\delta^{34}\text{S}$ in sulphates^{1,20}. These two values ($\delta^{13}\text{C}$ and $\delta^{34}\text{S}$) are negatively correlated if one uses mean values for periods. The model showing the relation between the carbon and sulphur cycle based on isotopic data, assuming scales for mixing in the sedimentary cycle of the order of 120–240 Myr, is given by Schidlowski *et al.*³.

Our understanding of the processes which cause the shift in isotopic ratios will be greatly improved by detailed studies of the transition zone in which the change in ¹³C is recorded. Magaritz *et al.*¹³ noted the presence of the transition zone at the base of the Zechstein sequence in the Marl Slate and its German equivalent the Kupferschiefer. We now discuss the results of a detailed investigation of this isotopic transition zone within the Marl Slate of North-East England.

The Marl Slate represents the earliest extensive marine transgression into the Rotliegend inland drainage basin whose floor lay well below the contemporaneous sea level in Permian times²¹. The preservation of unconsolidated dunes beneath the Marl Slate indicates that the Zechstein transgression was rapid, forming an inland sea up to 250 m deep in places²². The Marl Slate is usually less than 1-m thick and mainly comprises organic-rich, dolomitic laminites and closely resembles much younger sapropelic sediments in which lack of bioturbation due to the absence of benthic fauna is a characteristic feature. The present results come from a single core (VT8) drilled by the National Coal Board off the coast of North-East England east of Seaham Harbour (NZ 486 509). This core shows two cycles

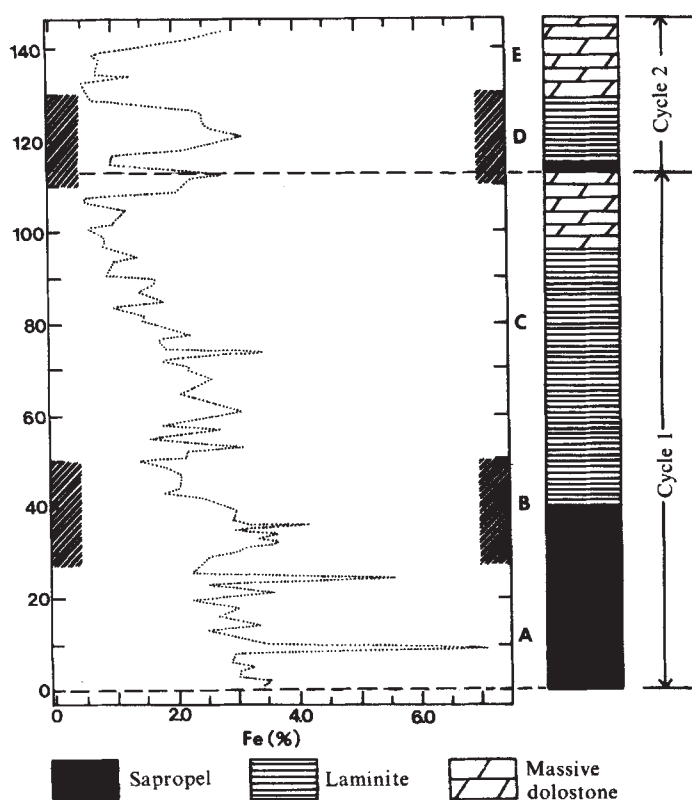


Fig. 2 Variation in iron content of the VT8 core. Note the similarity between the $\delta^{18}\text{O}$ curve and iron content.

of sedimentation which include the sequence: sapropel $\rightarrow$ dolomitic laminite $\rightarrow$ finely crystalline dolostone (Fig. 1). The core was divided into 144 1-cm thick segments which were used for isotopic and chemical measurements. The carbonate isotopic data (Fig. 1) indicate (1) the lack of correlation between $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ values; (2) a gradual enrichment in ^{13}C from the bottom (+2.3‰) to the top (+5.1‰); (3) a $\delta^{18}\text{O}$ pattern which exhibits spikes of very low values; (4) a very clear correspondence between the lithological variation and changes in $\delta^{18}\text{O}$, a feature which indicates that the isotopic changes are original and not the result of later recrystallization. The preservation of such detailed structure in the isotopic record of the Marl Slate is consistent with its deposition in an anoxic non-bioturbated environment. The organic-rich laminae are interpreted as seasonal phytoplankton blooms and thus enable assessment of the time scale of the events recorded in the core. Oelsner²³ counted the laminae in the Kupferschiefer and arrived at an age of 17,000 yr. Our observations are consistent with this suggesting an overall sedimentation of $\sim 0.08 \text{ mm yr}^{-1}$. This value is close to the rate estimated ($0.05\text{--}0.07 \text{ mm yr}^{-1}$) for the equivalent unit (the Stinkschiefer) in the second Zechstein cycle²⁴ but greater than the rate of sedimentation of anoxic sediments in the Lower Cretaceous North and South Atlantic ocean which is estimated²⁵ at 0.01 mm yr^{-1} . Our overall estimate of the time represented by the VT8 core is 18,000 yr and the large change in $\delta^{13}\text{C}$ values took place within this time interval. The core section can be divided into five macro-segments on the basis of the isotopic variation (Fig. 1), and the main carbon isotope shift occurs in the 22 cm of segment B. This is probably equivalent in time to $\sim 3,000 \text{ yr}$. Even shorter period events are detected in the core in both the oxygen and carbon isotopic record. The shift of more than 1‰ in $\delta^{13}\text{C}$ values which occurs within 3 cm ($\sim 400 \text{ yr}$) at the bottom of zone A is larger than any shift reported in the Cretaceous⁹. The $\delta^{18}\text{O}$ spikes, which are seen in several parts of the core occur within depositional periods of $< 5 \text{ cm}$, represent flooding of the basin by freshwater. The period of flooding is similar to that recorded in the Pleistocene sapropelic sediments of the

eastern Mediterranean²⁶. The low $\delta^{18}\text{O}$ values coincide with enrichment of iron (as pyrite) in the sediment (Fig. 2). This correlation indicates that the formation of anoxic conditions was associated with freshwater flooding of the basin. We take this to indicate that in these conditions there was an abundant supply of iron which enhanced iron sulphide formation.

The gradual change in $\delta^{13}\text{C}$ values in zone B and the early cycle of changes in zone A are considered to represent a change in the dissolved carbonate in the Zechstein Sea. From that point on the $\delta^{13}\text{C}$ values of the carbonates in the higher Zechstein cycles remain highly enriched in ^{13}C (refs 13, 14, 19). Such enrichment in ^{13}C is also seen in basins of Permian age other than the Zechstein in various parts of the world^{16–18}.

We argue that this change in the isotopic composition of the dissolved carbonate in the Zechstein Sea must reflect a similar change on an ocean wide scale. This is because a continuous supply of dissolved salts is needed to account for the succeeding thick carbonate/evaporite sequence. There is no obvious alternative supply of dissolved materials in view of the fact that the other source of dissolved carbonate (freshwater reflux) would result in low $\delta^{13}\text{C}$ values. This is evidenced by the less positive $\delta^{13}\text{C}$ values seen in the zones of $\delta^{18}\text{O}$ minima (Fig. 1).

It is not clear whether all the water column of the Permian ocean (all the oceanic bicarbonate reservoir) changed its carbon isotopic composition (reflecting a large variation in the C–S–N cycle) or whether the change occurred only in the top oceanic layer, because of the lack of a deep oceanic sequence for comparison. The size of the shift in $\delta^{13}\text{C}$ values in the Permian period is at least twice the size of any shift in the Cretaceous. Simple mass isotope calculations show that a shift of 3‰ in $\delta^{13}\text{C}$ will correspond to almost double the amount of organic material removed from the system at a specific time. Enrichment of 0.7‰ in $\delta^{13}\text{C}$ was estimated for the mean dissolved carbonate in the ocean water in a period $< 10,000 \text{ yr}$ for the Pleistocene glacial/interglacial transition²⁷. The shift recorded in the Permian probably took place at a similar rate, but the effect was tripled ($\sim 2.3\%$).

The $\delta^{34}\text{S}$ values of the Zechstein sulphates are all depleted in ^{34}S (ref. 28) which indicates that the accompanying effect in the sulphur cycle had already taken place by the end of the first Zechstein cycle. The very widespread regression ending the Permian Period and the Zechstein Sea was followed by the return to normal $\delta^{13}\text{C}$ values in early Triassic carbonates²⁹.

Received 20 November 1981; accepted 31 March 1982.

- Garrels, R. M. & Perry, F. A. in *The Sea* Vol. 5 (ed. Goldberg, E. D.) 303–336 (Wiley, New York, 1975).
- Mackenzie, F. T. & Pigott, J. D. *J. geol. Soc. Lond.* **138**, 183–196 (1981).
- Schidlowski, M., Junge, C. E. & Pietrek, H. *J. geophys. Res.* **82**, 2557–2565 (1977).
- Jeffery, P. M., Compston, W., Greenhalch, D. & Debacter, Y. *Geochim. cosmochim. Acta* **7**, 255–286 (1975).
- Weber, J. *Geochim. cosmochim. Acta* **31**, 2343–2351 (1967).
- Shackleton, N. J. in *The Fate of Fossil Fuel CO_2 in the Ocean* (eds Anderson, N. R. & Malahoff, A.) 543–563 (Plenum, New York, 1977).
- Fischer, A. G. & Arthur, M. A. *SEPM. Spec. Publ.* **25**, 19–50 (1977).
- Boersma, A. & Shackleton, N. J. *Cretaceous–Tertiary Boundary Event Symp.* **2**, 40–54 (1979).
- Scholle, P. A. & Arthur, M. A. *Bull. geol. Soc. Am.* **64**, 67–87 (1980).
- Kroopnick, P. M., Margolis, S. V. & Wang, C. S. in *The Fate of Fossil Fuel CO_2 in the Ocean* (eds Anderson, N. R. & Malahoff, A.) 295–321 (Plenum, New York, 1977).
- Jenkyns, H. C. *J. geol. Soc. Lond.* **137**, 171–181 (1980).
- Veizer, J., Holser, W. T. & Wilgus, K. K. *Geochim. cosmochim. Acta* **44**, 579–588 (1980).
- Magaritz, M., Turner, P. & Kading, K.-Ch. *Geol. J.* **16**, 243–254 (1981).
- Magaritz, M. & Schulze, K.-H. *Cont. Sedim.* **9**, 269–277 (1980).
- Ponov, A. B. *Geokhimiya* **8**, 1252–1277 (1976).
- Compston, W. *Geochim. cosmochim. Acta* **18**, 1–22 (1960).
- Dean, W. E., Davies, G. R. & Anderson, R. Y. *Geology* **3**, 367–372 (1975).
- Osaki, S. *Geochim. J.* **6**, 463–477 (1973).
- Botz, R. thesis, Univ. Heidelberg (1979).
- Thode, H. G., Monster, J. & Dunford, H. B. *Bull. Am. Ass. petrol. Geol.* **42**, 2619–2641 (1958).
- Smith, D. B. *J. geol. Soc. Lond.* **136**, 155–156 (1979).
- Smith, D. B. *Contr. Sedim.* **9**, (1980).
- Oelsner, O. *Freiberg. Forsch. C.* **58**, 106–113 (1959).
- Brongersma-Sanders, M. *Mar. Geol.* **11**, 123–144 (1971).
- Arthur, M. A. & Natland, J. H. in *Results of Deep Drilling in the Atlantic Ocean* Vol. 3, 375–401 (U.S. Govt. Printing Office, Washington DC, 1979).
- Ryan, W. B. F. & Cita, M. B. *Mar. Geol.* **23**, 197–215 (1977).
- Broecker, W. S. in *International Conference on Evolution of Planetary Atmospheres and Climatology of the Earth*, 165–190 (Cent. Nat. Etud. Spat. (France), 1978).
- Holser, W. T. & Kaplan, I. R. *Chem. Geol.* **1**, 93–135 (1966).
- Wilgus, C. K. thesis, Univ. Oregon, (1981).

Hotspots, polar wander, Mesozoic convection and the geoid

Don L. Anderson

Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA

The geoid bears little relation to present tectonic features of the Earth other than trenches and hotspots. The Mesozoic supercontinent of Pangea, however, apparently occupied a central position in the Atlantic–African geoid high. This and the equatorial Pacific geoid high contain most of the world's hotspots^{1,2}. The plateaus and rises which are now in the western Pacific formed in the Pacific geoid high and this may have been the early Mesozoic–late Palaeozoic position of a large part of Asia and other fragments of the Pacific rim continents. The major global geoid highs were regions of extensive Cretaceous volcanism and may be the former sites of continental aggregations and mantle insulation and, therefore, hotter-than-normal mantle. The pent-up heat causes rifts and hotspots and results in uplift, magmatism, fragmentation and dispersal of the continents and the subsequent formation of plateaus, aseismic ridges and seamount chains which cause a global rise in sea-level. Convection in the upper mantle caused by such lateral temperature gradients is intrinsically episodic. A geoid anomaly of 50 m can be formed in about 100 Myr by continental insulation. We show here that such geoid anomalies are long-lived and may be used to remove the ambiguity in early Mesozoic–late Palaeozoic plate reconstructions. Geoid highs control the rotation axis of the Earth and, in effect, bring long-lived continental aggregations to the Equator. Many aspects of continental geology such as vertical-tectonics and episodicity of magmatism and transgressions can be explained by continental insulation.

The Earth's largest positive geoid height anomalies are associated with subduction zones and hotspots and bear no simple relationship to other present-day tectonic regions such as continents and ridges. When the subduction-related geoid highs are removed from the observed field the residual geoid shows broad highs over the central Pacific and the eastern Atlantic–African regions^{1,2}. Like the total geoid, the residual geoid does not reflect the present-day distribution of continents and oceans and shows little trace of the ocean ridge system. Geoid highs, however, correlate with hotspots^{1,2} and with regions of anomalously shallow ocean floor and sites of extensive Cretaceous volcanism.

The Atlantic–African geoid high extends from Iceland through the North Atlantic and Africa to the Kerguelen plateau and from the middle of the Atlantic to the Arabian peninsula and western Europe. Most of the Atlantic, Indian Ocean, African and European hotspots are inside this anomaly. Iceland, Trindade, Tristan, Kerguelen, Reunion, Afar, Eiffel and Jan Mayen form the 20-m boundary of the anomaly and seem to control its shape. The Azores, Canaries, New England seamounts, St Helena, Crozet and the African hotspots are interior to the anomaly. This, plus similar evidence in the Pacific, suggests that geoid highs may be associated with hotter-than-normal mantle.

Although the geoid high cuts across present-day ridges and continents there is a remarkable correspondence of the predrift assemblage of continents with the present location of both the geoid anomaly and hotspots (Fig. 1). Reconstruction of the mid-Mesozoic configuration of the continents³ reveals, in addition, that most of the large shield areas of the world are contained inside the Atlantic–African geoid high. Most of the Phanerozoic platforms are also in this area, the main exceptions being the Siberian and South-East Asian platforms.

The area inside the geoid high is also characterized by higher-than-normal present elevations, for example, in Africa, the

North Atlantic⁴ and the Indian Ocean south-east of Africa⁴. This also holds true for the axial depth of oceanic ridges⁵. Most of the continental areas were above sea level from the Permian–Carboniferous to the Triassic, at which time there was subsidence in eastern North and South America, central and southern Africa, Europe and Arabia⁶. The widespread uplift, magmatism, breakup and initial dispersal of the Pangean land-mass apparently occurred while the continents were centrally located with respect to the present geoid anomaly. The subsequent motions of the Atlantic bordering plates were largely directed away from the anomaly. This suggests that the residual geoid high, hotspots, the distribution of continents during the late Palaeozoic and early Mesozoic and their uplift and subsequent dispersal and subsidence are all related. The shields are regions of abnormally thick lithosphere. The thickest lithosphere is in eastern and central North America, northeastern South America, northwestern and central Africa and northern Siberia⁷. These regions were all within the area of Fig. 1 at 200 Myr.

The area in Fig. 1 also experienced exceptional magmatism during the Mesozoic. The great flood basalts of Siberia and South Africa were formed during the Triassic and Jurassic, possibly at the sites of the Jan Mayen and Crozet hotspots³. The plateau basalt provinces of South-east Greenland and Brazil were formed during the Cretaceous and early Cenozoic, possibly at the sites of the Iceland and Tristan hotspots³. The Walvis Ridge, the Rio Grande Rise and the plateaus in the western Indian Ocean are mainly on Mesozoic and Palaeocene crust. A large part of the Pacific also experienced extensive on- and off-ridge volcanism in the Cretaceous^{8,9}. This extensive ridge and hotspot volcanism may have caused the rapid rise in sea level during the Jurassic and Cretaceous⁹. If sea level¹⁰ can be used as a guide to the volume of the ocean basins then from the end of the Cretaceous to the end of the Oligocene was a period of less intense oceanic volcanism and subsidence of the oceanic and continental crust. Sea-level variations may therefore indicate that the thermal and geoid anomalies formed in the Palaeozoic and have attenuated since the early Mesozoic.

Much of the Pacific rim appears to be terrain which originated in the Pacific^{11,12}. The possibility of a continent centrally located in the Pacific has been discussed for some time¹³, but its location has been an enigma and its size uncertain. The central Pacific geoid anomaly, like the Atlantic–African anomaly, may mark the early Mesozoic or late Palaeozoic location of a large continental landmass and, later, the site of extensive Cretaceous ridge crest and midplate volcanism.

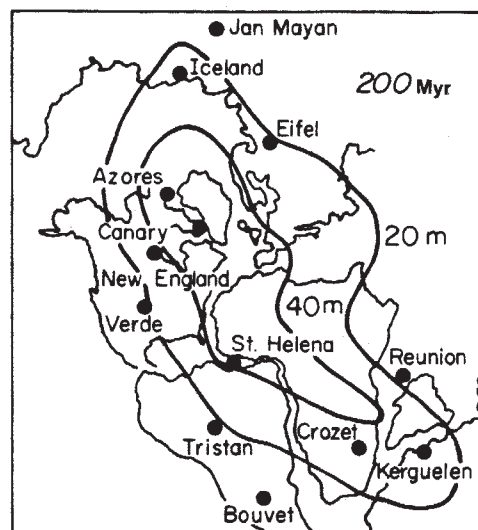


Fig. 1 The Jurassic configuration of the continents, based on Morgan's hotspot-based reconstruction³, superposed on the Atlantic–African residual geoid high^{1,2}. Note that most of the stable shield areas are inside the 20-m contour. Selected hotspots (●) are also shown.

Palaeomagnetic data indicate that various blocks of Asia such as Kolyma, Sikhote Alin, Sino-Korea, Yangtze, South-east Asia and Japan have moved northwards by up to 32° since the Permian¹¹. It is unlikely that they were in the vicinity of Australia or associated with Gondwana¹¹ and a central Pacific location is likely. Part or all of Alaska and northwestern North America were also far south of their present positions, relative to North America, in the Permian¹². The same may be true of California, Mexico, Central America and other accreted terrain in the Pacific rim continents.

The central Pacific residual geoid high (>20 m) extends from Australia to the East Pacific Rise and from Hawaii to New Zealand. It encompasses most of the Pacific hotspots and is approximately antipodal to the Atlantic-African anomaly.

The western Pacific contains numerous plateaus, ridges, rises and seamounts which have been carried far to the north-west from their point of origin. The Ontong-Java plateau for example, presently on the Equator, was formed at 33–40°S in the mid-Cretaceous¹⁴. The Hess rise, Line Islands ridge and Necker ridge were formed in the Cretaceous on the Pacific, Farallon and Phoenix ridge-crests⁸ which were, at the time, in the eastern Pacific in the vicinity of the polynesian seamount province. This ridge-crest volcanism was accompanied by extensive deep-water volcanism^{8,9} and, possibly, rapid seafloor spreading¹⁵. The Caribbean and Bering Seas may have formed at the same location and carried northward on the Farallon and Kula plates. It is significant that the Caribbean and the anomalous regions in the Pacific have similar geophysical and geochemical characteristics¹⁶.

The Pacific geoid high also has anomalously shallow bathymetry^{4,5}. This shallow bathymetry, the 'Darwin Rise', extends from the East Pacific Rise to the north-west Pacific in the direction of plate motion, and includes the central and south-west Pacific hotspots. The extensive volcanism in the central western Pacific between ~70 and 120 Myr (refs 8, 9, 16) occurred about 60°–100° to the south-east of its present position, in the hotspot frame. This would place the event in the vicinity of the southeastern Pacific hotspots and the eastern part of the geoid anomaly. This suggests that the anomaly dates back to at least early Cretaceous. A similar thermal event in the Caribbean¹⁷ may mean that it was also in this region in the Cretaceous, particularly since the basalts in the Caribbean are similar to those in the western Pacific¹⁶.

Going back even further are the Triassic basalts in Wrangellia¹¹ of northwestern North America and the Permian greenstones in Japan¹³, both of which formed in equatorial latitudes and subsequently drifted to the north and north-west respectively. We suggest that they also formed in the Pacific geoid high.

Plate reconstructions based on palaeomagnetism or palaeoclimatology have arbitrary longitudinal relationships between the plates. Figure 2 gives a possible late Palaeozoic reconstruction which places Africa, South America, Antarctica and Europe in the Atlantic-African geoid high and Asia in the Central Pacific geoid high. This is similar to previous reconstructions^{6,18} in that North America is adjacent to western South America but Asia is placed much closer to North America. The biogeographic similarities between Asia and western North America are well known¹⁹ and the placement of North America provides the continuity between the major land masses which is suggested by some palaeontological data. With this reconstruction Gondwanaland and Asia were the continental aggregations that caused the geoid anomalies and they brought themselves to the Equator, by the Goldreich-Toomre mechanism²⁰, before their breakup and dispersal. Parts of Alaska, western North America and West Antarctica may also have migrated away from the Pacific geoid high.

We propose that episodicity in continental drift, polar wander, sea-level variations and magmatism are due to the effect of thick continental lithosphere on convection in the mantle and that supercontinents 'insulated' large parts of the mantle for more than 10⁸ yr. The excess heat caused uplift by thermal expansion and partial melting and, eventually, breakup and

dispersal of the supercontinents. A large geoid high is generated by this expansion and, if this is the dominant feature of the geoid, the spin axis of the Earth would change so as to centre the high on the Equator²⁰, much as it is today. The continents stand high while they are within the thermal and geoid high and subside as they drift over cooler mantle.

The association of the Atlantic-Africa geoid high with the former position of the continents and the plateau basalt provinces with currently active hotspots suggests that Mesozoic convection patterns in the mantle still exist, as proposed by Menard²¹. The mid-Atlantic and Atlantic-Indian ridges and the Atlantic, African and Indian ocean hotspots and the East African rift are regions where heat and material is being efficiently removed from the uppermantle in the Atlantic-African geoid high. The East Pacific rise and the Pacific hotspots are serving the same role in the Pacific geoid high. If geoid anomalies and hotspots are due to a long period of continental insulation and if these regions had higher-than-normal heat flow and magmatism for the past 100 Myr, then we might expect that these features will wane with time.

Global isostatic geoid anomalies can be written^{22,23}

$$N \cong \frac{2\pi G}{g} \frac{(l+2)}{(2l+1)} \rho h \delta h$$

where G is the gravitational constant, g is the acceleration of gravity, l is degree of spherical harmonic, ρ is density, h is the depth of the anomalous region and δh is the elevation anomaly associated with the mass anomaly. For a coefficient of thermal expansion of $3 \times 10^{-5} \text{ } ^\circ\text{C}^{-1}$, a temperature excess of 70 °C in the upper 500 km of the mantle gives $\delta h = 1 \text{ km}$ and $N = 50 \text{ m}$ for $l = 3$. A similar effect can be achieved by 1% melting. Some combination of heating and melting is likely.

The average heat loss through oceanic areas is $2.4 \text{ } \mu\text{cal cm}^{-2} \text{ s}^{-1}$ (ref. 24). Continental heat flux is about 0.6 of this in spite of the higher radioactivity of continental crust. If we assume that the oceanic heat flux is available everywhere and that there is a balance between heat production and heat loss under oceans the temperature rise under an extended stationary continent is $0.60 \text{ } ^\circ\text{C Myr}^{-1}$ for a 500-km thick column or, alternatively, 6% melting, using a latent heat of 100 cal g^{-1} . Thus, 100 Myr of continental insulation can give the observed geoid and elevation anomalies. The continental aggregations were relatively stable for at least this long^{6,18}. True polar wander is likely while the geoid anomalies are developing.

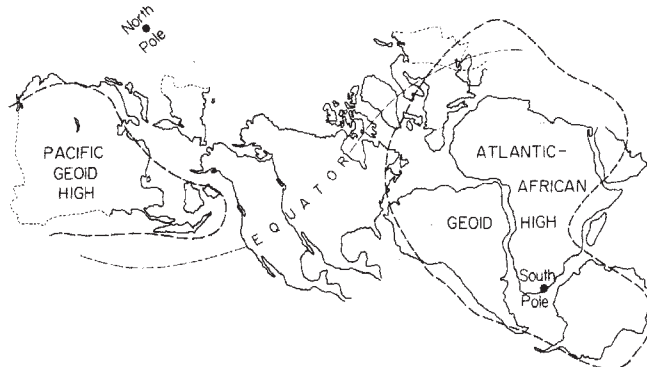


Fig. 2 Possible late Palaeozoic continental configurations relative to the current residual geoid highs and the Palaeoequator. This is similar to previous reconstructions^{6,18} except that the Pacific geoid high is used to position eastern Asia. Australia and India are not shown. The dashed lines enclose regions of high elevation (>20 m) of the residual geoid^{1,2}. If the geoid highs were generated under Gondwana and Asia they formed in high latitudes and were subsequently brought to the Equator by their affect on the rotation axis of the Earth. The geoid highs are currently centred on the Equator. Parts of western North America may also have been in the Pacific geoid high. This and the eastern migration of North America, relative to South America, during the Palaeozoic are indicated schematically by plotting North America in two positions. Asia is probably also fragmented at this time¹¹.

After the continents move off the mantle which they have been insulating the thermal and geoid anomalies decay by convection in the mantle and conduction through the newly-formed oceanic lithosphere. The rate of temperature decrease, $\dot{\theta}$, can be written

$$\dot{\theta} = \frac{-\kappa\theta}{h^2} \left(\frac{R}{R_{cr}} \right)^{1/3}$$

where κ is the thermal diffusivity, R is the Rayleigh number and R_{cr} is the critical Rayleigh number. For $R_{cr} = 10^3$, $R = 10^6$, $\kappa = 10^{-2} \text{ cm}^2 \text{ s}^{-1}$, $\theta = 10^3 \text{ K}$ and $h = 500 \text{ km}$

$$\dot{\theta} = -0.12^\circ \text{C Myr}^{-1}$$

This gives a decrease in elevation of 180 m in 100 Myr. The corresponding geoid height reduction is $\sim 9 \text{ m}$. Thus, geoid anomalies caused by continental insulation are long-lived features.

Horizontal temperature gradients can drive continental drift²⁵⁻²⁸. The velocities decrease as the distance increases away from the heat source and as the thermal anomaly decays. The geoid high also decays. Thick continental lithosphere then insulates a new part of the mantle and the cycle repeats. Periods of rapid polar motion and continental drift follow periods of continental stability and mantle insulation.

The relatively slow motion of the continents²⁹ or the pole³⁰, or both, during the Permian and the relative stability of sea level during this period suggest that the geoid anomalies were developing at this time. Continental drift velocities²⁹ and sea level¹⁰ changed rapidly during the Triassic and Jurassic. This we interpret as the start of extensive rift and hotspot magmatism and the decay of the thermal and geoid anomalies.

The rotation axis was apparently stable from ~ 200 to 80 Myr (ref. 3) but has shifted with respect to the hotspot reference frame for at least the past 65 Myr (ref. 31). In the framework of our model this shift would represent a growth of the equatorial region of the Atlantic-African high due to continued insulation by Africa, a decay of the South Pacific or North Atlantic portions of the highs due to extensive Cretaceous volcanism, a reconfiguration of the world's subduction zones or some combination.

It is difficult to separate true polar wander and the apparent wander due to continental drift. In the present model they are intimately related. If ridge volcanism is precluded for an extensive period of time, as seems to have been the case in the mantle under Pangea during the early Mesozoic and late Palaeozoic, the mantle will warm up and/or partially melt. In either case a geoid anomaly will develop. This anomaly will affect convection and plate motions in a manner previously suggested^{25,32} and will also affect the rotation axis of the Earth²⁰. As the thermal anomalies seem to be long-lived and it takes a long time to establish a new thermal regime, major shifts of the rotation axis due to this mechanism will be infrequent and gradual.

Goldreich and Toomre²⁰ have shown that density inhomogeneities caused by convection in the mantle will control the orientation of the spin axis and that large shifts of the pole, that is, true polar wander, are to be expected in a dynamic Earth. The present orientation of the pole apparently bears no relationship to the present distribution of continents, ridges or trenches³². On the other hand, the geoid bears a close relationship to the present distribution of hotspots and to the former distribution of continents and areas of extensive Cretaceous ridge and plateau volcanism.

If our conjecture about the relationship between geoid highs and the former position of continents is correct, then we would conclude that there were two large antipodal continental aggregations in the Permian. The antipodal supercontinents were presumably formed by a previous episode of continental drift which swept buoyant material away from an earlier configuration of thermal and geoid highs. Using sea level as a guide, with high sea-level indicating rifting and extensive sea floor volcanism, the supercontinents were stable from 300 to 150 Myr

ago and the previous continental assemblages were dispersing between 500 and 350 Myr. The suggestions that hotspots may be a primary cause of continental breakup³ and that they affect polar wander^{1,2} are not new. What is new is the suggestion that continents cause geoid anomalies and hotspots and, indirectly, control the rotation axis of the Earth.

I thank Brad Hager, Clem Chase, Kevin Burke and Henri Nataf for reviewing the manuscript, and Lynn Sykes for stimulating discussions. Art Boucot gave useful advice on biogeography. Brad Hager showed me how to calculate global isostatic geoid anomalies. This work was supported by NASA grant NSG-7610. Contribution no. 3710, Division of Geological and Planetary Sciences, California Institute of Technology.

Received 16 November 1981; accepted 27 January 1982.

- Chase, C. *Nature* **282**, 464-468 (1979).
- Crough, S. & Jurdy, D. *Earth planet. Sci. Lett.* **48**, 15-22 (1980).
- Morgan, W. J. in *The Sea* Vol. 7 (ed. Emiliani, C.) 443-488 (Wiley-Interscience, New York 1981).
- Cochran, U. & Talwani, M. *Geophys. J.* **50**, 495-552 (1977).
- Menard, H. & Dorman, L. *J. geophys. Res.* **82**, 5329-5335 (1977).
- Kanasewich, E. R., Havskov, J. & Evans, M. *Can. J. Earth Sci.* **15**, 919-955 (1978).
- Pollack, H. & Chapman, D. S. *Tectonophysics* **38**, 279-296 (1977).
- Watts, A. B., Bodine, V. & Ribe, N. *Nature* **283**, 532-537 (1980).
- Schlanger, S. O., Jenkyns, H. C. & Premoli-Silva, I. *Earth planet. Sci. Lett.* **52**, 435-449 (1981).
- Vail, P., Mitchum, R. M. Jr & Thompson, S. III *Am. Ass. petrol. Geol. Mem.* **26**, 83-97 (1978).
- McElhinny, M., Embleton, B. J. J., Ma, X. H. & Zhang, Z. K. *Nature* **293**, 212-216 (1981).
- Panaska, B. & Stone, D. *Nature* **293**, 561-563 (1981).
- Hattori, I. & Hirooka, K. *Tectonophysics* **57**, 211-235 (1979).
- Hammond, S., Kroenke, L., Thayer, F. & Keelson, D. *Nature* **255**, 46-47 (1975).
- Larson, R. & Chase, C. *Bull. geol. Soc. Am.* **83**, 3627-3644 (1972).
- Batiza, R., Larson, R., Schlanger, S., Scheka, S. & Tokuyama, H. *Nature* **286**, 476-478 (1980).
- Burke, K., Fox, D. & Senor, A. M. *J. geophys. Res.* **83**, 3949-3954 (1978).
- Smith, A. G., Hurley, A. M. & Briden, J. C. *Phanerozoic Paleogeographic World Maps* (Cambridge University Press, 1981).
- Boucot, A. J. & Gray, J. in *Historical Biogeography, Plate Tectonics, and the Changing Environment* (eds Gray, J. & Boucot, A. J.) 465-482 (Oregon State University Press, 1979).
- Goldreich, P. & Toomre, A. *J. geophys. Res.* **74**, 2555-2567 (1969).
- Menard, H. W. *Earth Planet. Sci. Lett.* **20**, 237-241 (1973).
- Dahlen, F. A. (in preparation).
- Hager, B. H. (in preparation).
- Slater, J. G., Parsons, B. & Jaupart, C. *J. geophys. Res.* **86**, 11,535-11,552 (1981).
- Elder, J. *Nature* **214**, 657-660 (1967).
- Ichije, T. *J. geophys. Res.* **76**, 1139-1153 (1971).
- Busse, F. J. *Geophys. J. R. astr. Soc.* **52**, 1-12 (1978).
- Froidevaux, C. & Nataf, H. *Sonderdruck aus der Geologischen Rundschau Band 70*, 166-176 (1981).
- Gordon, R. G., McWilliams, M. O. & Cox, A. J. *J. geophys. Res.* **84**, 5480-5486 (1979).
- Van Alstein, D. R. thesis, California Inst. Technol. (1979).
- Gordon, R. & Cape, C. *Earth planet. Sci. Lett.* **55**, 37-47 (1981).
- Jurdy, D. J. *J. geophys. Res.* **83**, 4989-4994 (1978).

Development of earthquake-induced fissures in the Main Ethiopian Rift

Laike M. Asfaw

Geophysical Observatory, Addis Ababa University, PO Box 1176, Addis Ababa, Ethiopia

A series of episodes involving land subsidence along an elongated formation of pits connected by fissures have been observed in various parts of the Main Ethiopian Rift¹⁻⁴ over the past 25 yr. Although other workers¹ have proposed a possible relationship between tectonics, subsurface drainage and subsidence no field observations have related earthquakes to such development of fissures. I present here the first observation of the development of fissures following an earthquake swarm with local magnitude, $M_L \leq 4$ in the Main Ethiopian Rift north of the Fentale Mountain (Fig. 1) between January and March 1981 (ref. 5). This showed that there could be a variety of developments in the evolution of earthquake-induced fissures depending on geological formation, vegetation cover and drainage; the implication on similar earlier observations is considered.

That the Afar Depression and the Main Ethiopian Rift are regions under tension is well established through geodetic

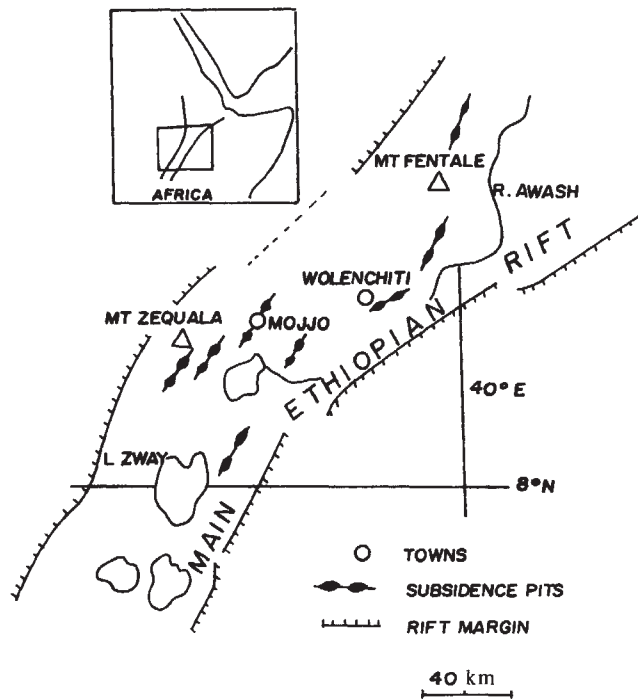


Fig. 1 Sites of subsidence pits associated with fissures in the Main Ethiopian Rift including the event of January–March 1981 north of the Fentale Mountain.

measurements and other surface observations^{6,7}. The surface observations consist of tensional fissures and faults along the three major rift trends of the Red Sea, Gulf of Aden and the Main Ethiopian Rift^{8,9}. In the Main Ethiopian Rift the Wonji Fault Belt is characterized by fissures and faults¹⁰ and sometimes fissures can be identified by elongated green belts of vegetation in this rift. Here the sediment deposit, rainfall and vegetation make the development of fissures more complicated and varied compared with those in the more arid rifts in the Afar Depression trending in the Red Sea and the Gulf of Aden directions.

Earlier investigators^{1–4} had noted subsidence and erosion associated with fissures in the Main Ethiopian Rift. In most of the cases the features observed consist of subsidence pits connected by fissures and oriented either obliquely or parallel to the rift trend, generally indicating a relation to the tectonics of the region. Gouin and Mohr¹ tried to relate earthquakes which occurred a few months earlier in 1966 with such subsidence in the rift floor; they suggested a possible relationship between tectonics, underground drainage and surface subsidence, however, neither a direct relation to earthquakes nor the actual development of the fissures was observed. All reports of subsidences associated with these type of formations were made after the incidences, usually following heavy rains, and therefore the problem of their origin was not solved.

The occurrence of fissures during an earthquake swarm in the Main Ethiopian Rift north of the volcanic centre of Fentale between January and March 1981 and the subsequent erosion provided the first opportunity to follow the development and identify the main features in such a process. In the epicentral area fissures, 2 cm wide, run on sediments and basalt flows for a few kilometres and had a NNE–SSW orientation following the rift trend. After an interval of 6 months, which included a rainy season, the fissures that occurred on sediments on the site were found to have changed considerably while fissures on basalt flows showed very little changes.

Fissures in the sedimentary layer became surface and subsurface drainage channels and the features that developed during a subsequent erosion were determined by the characteristics of sites such as the vegetation cover and the exact nature of

geological formations. Depending on the vegetation cover the erosion of fissures on sedimentary formation was either subsurface, where there was deep-rooted vegetation, or gully type, where there was no substantial growth (see Figs 2b and 3c). In the case of subsurface erosion of fissures the subsidence of the thin sedimentary veneer, which was held together by roots of vegetation, following heavy rains, resulted in the sporadic occurrence of pits connected by fissures (see Figs 2c and 3b). On the other hand, fissures on sediments with flat topography and therefore with minimum surface drainage were sealed off and did not develop further (see Figs 2a and 3a). The observed states of erosion following the earthquake swarm between January and March can be classified into three main categories as indicated in Fig. 2.

The features in the third type of development, Fig. 2c, bear a strong resemblance to many other observations of subsidences in the Main Ethiopian Rift. The depths of such formations were much larger than their widths and at times full-grown trees had been swallowed in the pits. The great depths attained compared with their widths can be explained by erosion due to subsurface drainage which would mainly be conditioned by the water table there. In such situations both the bottom and lower walls of the fissures would be eroded in an unrelated manner to the overlying walls and cover thereby deepening and widening the lower part disproportionately. Usually the pits occur where there is deep-rooted vegetation. Here the roots would fracture the sediment, facilitating the process of erosion underneath—especially laterally—widening the original fissure into pits. Of course, vegetation would minimize surface erosion and not necessarily subsurface erosion. The sporadic occurrence of pits is partly due to the sporadic nature of the distribution of deep-rooted vegetation and partly due to geological factors.

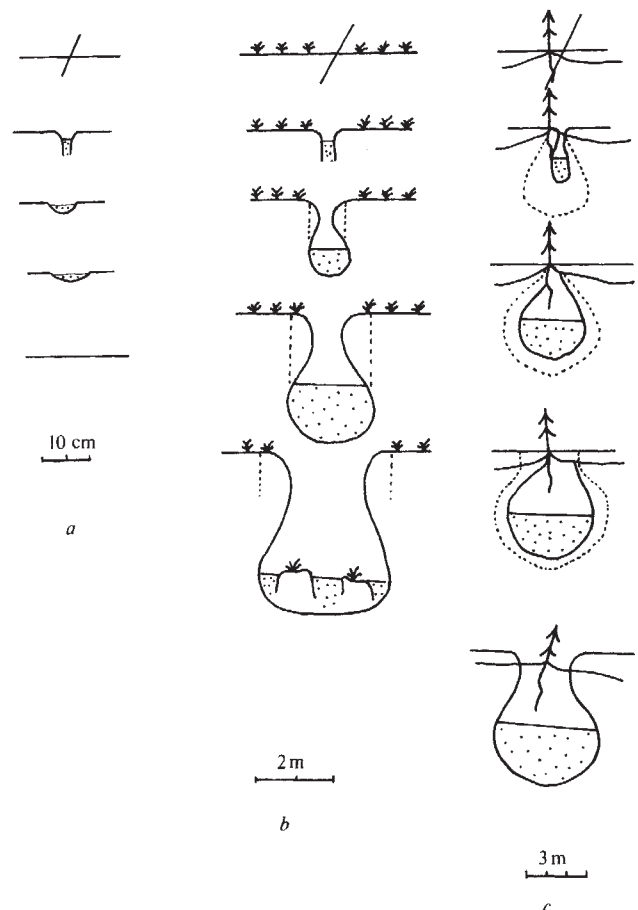


Fig. 2 Types of development of fissures on sediments: a, fissures on flat ground with poor drainage; b, gully type erosion of fissures; c, development of subsidence pits. Approximate scales do not include vegetation. No vertical exaggeration.

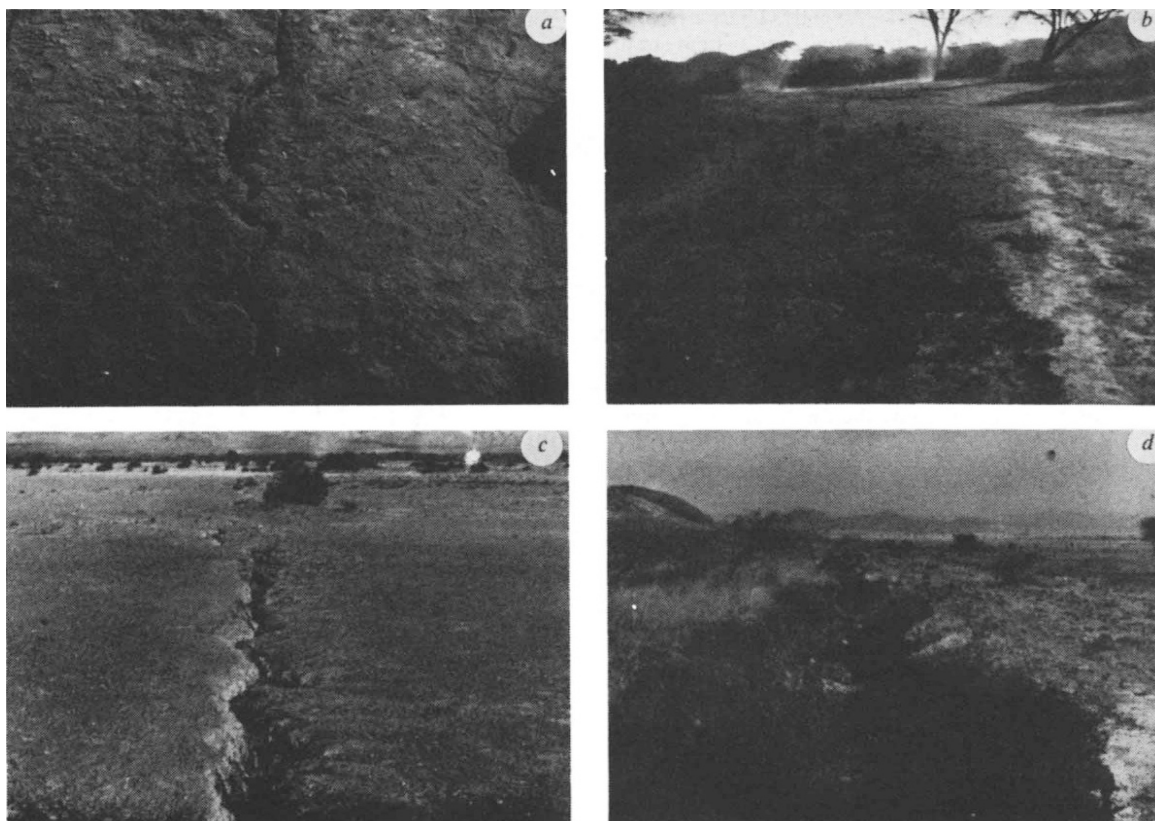


Fig. 3 *a*, Segment of a crack, 2-cm wide, which occurred during the earthquake swarm of January–March 1981 north of Fentale Mountain. This crack was sealed off 6 months later in June 1981. *b*, Subsequent development of pits, about 1 m wide and 50 cm deep, 6 months later, 500 m to the north of Fig. 3*a*. *c*, A small gully type development, 15 cm wide and 10 cm deep, 300 m to the south of Fig. 3*a*. *d*, Gully type erosion of a fissure, 2 m wide and 1.5 m deep, to the west of Fentale Mountain.

The orientation of pits and fissures relative to the rift trend is indicative of their tectonic origin where tension in the region would be associated with fissures oriented generally in the direction of the rift.

The third type of development of fissures, Fig. 3*b*, following the earthquake swarm of 1981, took only 6 months but the main features that characterize subsidence fissures elsewhere had developed. For example, the Mojjo and Wolenchiti fissures and subsidence pits^{1–3} could have been considerably older and their rates of development different due to variation in precipitation, subsurface drainage and geological formation. Allowing for age differences the third type of development fully accounts for all the features observed in various parts of the Main Ethiopian Rift. Among other things, the earthquake incidence between January and March 1981 showed that it is possible for earthquake swarms with local magnitude, $M_L \leq 4$ to be associated with fissures. This association is plausible in view of the thin crust there where seismic moments and therefore magnitudes of earthquakes would be small.

On the other hand the segment of the crack shown in Fig. 3*a* was sealed off 6 months later and no trace could be found, while 300 m to the south (Fig. 3*c*) a small gully type development was initiated. These developments correspond to the schemes shown in Figs 2*a* and *b* respectively.

Eruption at Le Piton de la Fournaise volcano on 3 February 1981

P. Bachelery, P. A. Blum, J. L. Cheminee, L. Chevallier, R. Gaulon*, N. Girardin*, C. Jaupart†, F. Lalanne, J. L. Le Mouel, J. C. Ruegg* & P. Vincent

Observatoires Volcanologiques, Institut de Physique du Globe de Paris, Tour 14–24, Université Paris VI, 4 Place Jussieu, 75230 Paris Cedex 05, France

Piton de la Fournaise in the island of La Réunion, Indian Ocean, is a young (0.35-Myr old) basaltic shield volcano built on the flank of the extinct Piton des Neiges volcano. It is one of the most active volcanoes in the world (magma output rate of $0.32 \text{ m}^3 \text{ s}^{-1}$)^{1,2} and has been under continuous surveillance by the Volcanological Observatory of the Institut de Physique du Globe de Paris since August 1980. The first detailed observations on the volcano reveal a remarkably low level of background seismicity. The 3 February 1981 eruption was preceded by a deep event of volcanic origin at depths of 10–15 km on 2 January; 20 days of quiescence followed this event, until the onset of shallow seismicity close to the eruptive centres on 22 January. The eruption began after 13 days of sustained shallow seismicity and occurred in three phases over 92 days and produced about $10 \times 10^6 \text{ m}^3$ of lava.

A preliminary report of the activity preceding the eruption is given in ref. 3. The seismic network consists of four vertical seismometers and one three-component seismometer (PHR) located around the caldera (Fig. 1*a*). The instruments have a linear response at frequencies of 1–30 Hz with a gain of 10^3 . Signals are transmitted by radio to the observatory where they

* Also at Laboratoire L.A. 195.

† To whom reprint orders should be addressed.

Received 20 October 1981; accepted 20 March 1982.

1. Gouin, P. & Mohr, P. *Bull. geophys. Obs.* **10**, 39–46 (1967).
2. *Smithsonian Instn. CFSLP Event Card No. 1019* (1970).
3. *Smithsonian Instn. CFSLP Event Card No. 1231* (1971).
4. Williams, M. A. J. *Sinet Ethiop. J. Sci.* **3**, 2 (1981).
5. Laike, M. A. *Proc. Conf. Earthquake Swarm (ECA in the press)*.
6. Tarantola, A. *et al. Earth planet. Sci. Lett.* **45**, 435–444 (1979).
7. Mohr, P. *et al. Tectonophysics* **44**, 141–160 (1978).
8. Varet, J. & Gasse, F. *Geology of Central and Southern Afar* (CNRS, Paris, 1978).
9. Harrison, C. G. A. *et al. in Afar Depression of Ethiopia*, 178–198 (Schweizerbart, Stuttgart, 1975).
10. Mohr, P. *Bull. geophys. Obs.* **11**, 1–65 (1967).

are recorded on strip-chart (2 mm s^{-1} speed). The number of stations was doubled and a new digital recording system set up in January 1982.

Magnitudes were computed using the relationship $M = 2.0 \log \tau - 0.9$, where τ is the event duration. The detection limit was $M = -1$ and several events of that magnitude may have been missed. Hypocentres were determined using a homogeneous model with a fixed V_P/V_S ratio of 1.7, and by varying the value of V_P between 4 and 7 km s^{-1} . The inter-station distance was typically 5 km and epicentres located in the network interior are considered accurate to 1 km. For those events which were shallow, errors in depth estimates are $\sim 2 \text{ km}$. For epicentres located outside the network, the errors are larger but $< 3 \text{ km}$ for the events reported here. The corresponding errors on hypocentre depths are probably as high as 5 km. A qualitative check on the locations is provided by the distribution of amplitudes.

The Piton de la Fournaise volcano is characterized by frequent volcanic activity with eruptions following each other at 1–2 yr intervals. Yet the seismicity of the island is remarkably low. Researches on local archives dating from 1800 have revealed no accounts of earthquakes of sufficient magnitude to be felt by the inhabitants outside the immediate vicinity of the volcano. During the 5 months which preceded the eruption, < 12 events per month were detected by the network (Fig. 2). Events of magnitude > 1 seldom occur and the magnitude 2 level was never exceeded. This is in sharp contrast with the level of background seismicity observed at Kilauea in Hawaii, a volcano which belongs to the same category (intra-oceanic).

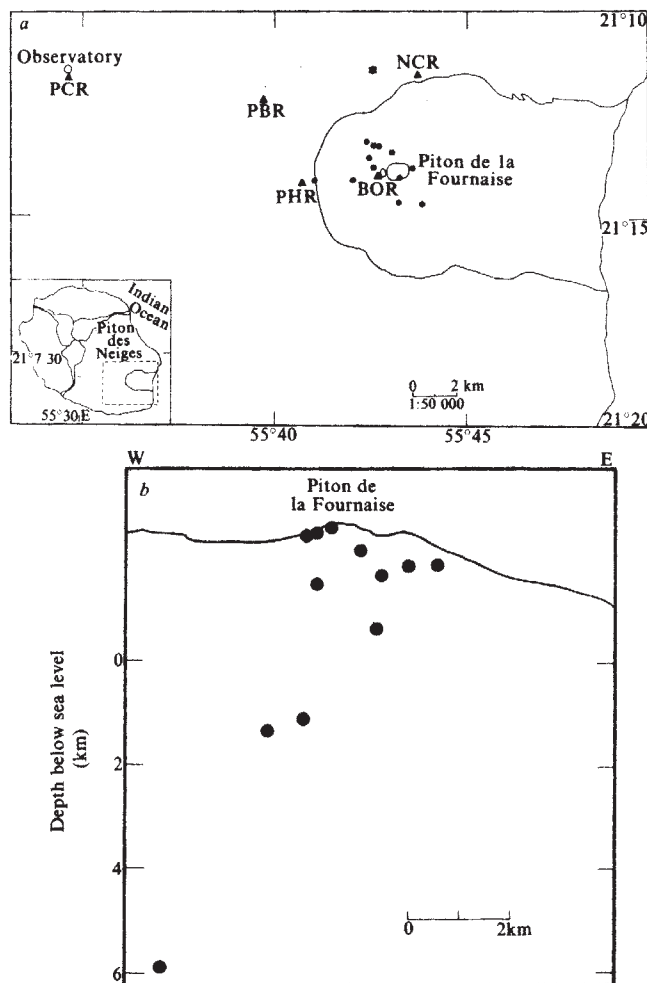


Fig. 1 *a*, Map of the volcano showing the seismic stations (Δ) and epicentres of selected events ($\bullet$). $\star$, The epicentre of the deep event of 2 January 1981. Locations are accurate to $\sim 1 \text{ km}$ in the interior of the network. *b*, Depths of the events shown on *a*. The error is about $\pm 2 \text{ km}$. The deep event at a depth of $\sim 6 \text{ km}$ occurred a few hours before the eruption start on 3 February 1981.

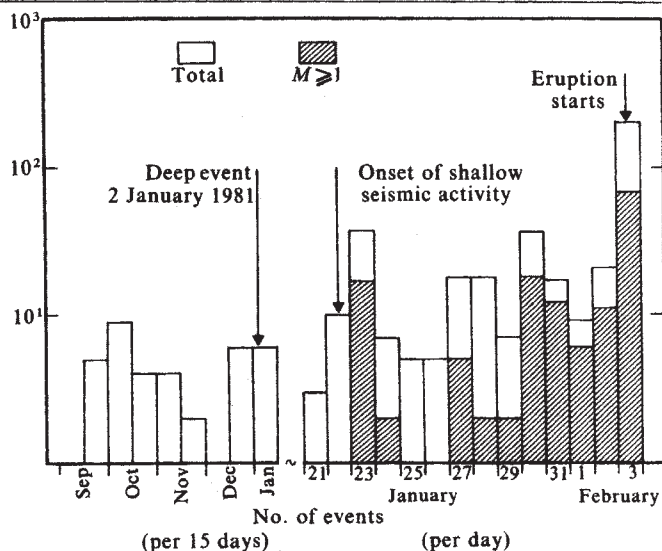


Fig. 2 Seismicity of the Piton de la Fournaise volcano. Note the change of scale on 21 January 1981.

Reports from the Hawaiian Volcano Observatory indicate that the monthly number of events with magnitudes > 1 and > 3 typically exceed 100 and 10 respectively (see, for example, refs 4, 5).

Premonitory signs of the volcano awakening were first detected on 2 January 1981, when a most unusual event was recorded. It lasted 5 min and was clearly not of tectonic origin (Fig. 3). The S-phase was followed by low frequencies below 0.5 Hz and P-wave polarities varied from station to station. It was a deep event with small-amplitude variations across the network, located 10–15 km below the northern part of the caldera (Fig. 1*a*). It was probably caused by the movement of a light phase (magma or gas) and not by fracturing. No similar event had been observed before and none has occurred since.

No significant increase of seismic activity was noted afterwards (Fig. 2) until 22 January 1981, when 10 events were observed on a single day. On the following day there were 38 events and a warning was issued to the local authorities on 24 January. Seismicity remained at a high level and dramatically increased just before the eruption (Fig. 2). This activity was located around the small Bory crater, close to station BOR (Fig. 1*a*), at depths smaller than 3 km (Fig. 1*b*), with the exception of a deeper event which occurred a few hours before the eruption.

A geodetic network of eight stations was set up around the volcano and the first measurements were carried out on 12 December 1980. The precision in the distance estimates is $\sim 3 \text{ cm}$. Following the onset of seismic activity on 22 January, measurements were conducted on 24 January and 30 January. Only the latter showed a small difference with those of 12 December 1980, barely above the threshold of error, which were interpreted as inflation of the Bory zone³. Subsequent measurements made after the eruption start on 12 February 1981 revealed a displacement of $18 \pm 2 \text{ cm}$ perpendicular to the system of fissures which had opened³. No further variations could be detected during February.

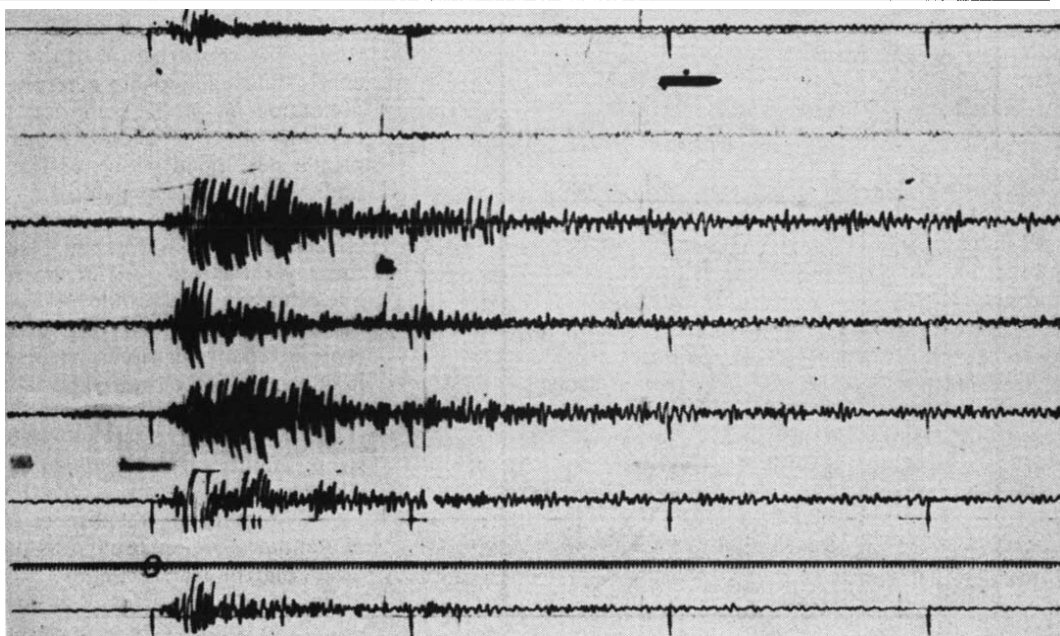
Unfortunately no measurements could be made before 12 December 1980. These would have allowed us to see whether

Table 1 Characteristics of the eruptive phases

Eruptive phase	Surface of lava flows (10^6 m^2)	Volume* (10^6 m^3)
1. 3–24 February	0.5	2.5
2. 24 February–31 March	0.7	3.5
3. 1 April–5 May	0.9	4.5

* Computed using a mean thickness of 5 m.

Fig. 3 The 2 January 1981 event. Large ticks on the records mark 1 min intervals. This event was located at a depth of 10–15 km (from ref. 3).



the inflation detected on 30 January had been going on for a longer time, before the deep volcanic event of 2 January. If such an inflation did occur, however, it was not associated with any significant seismic activity, which would be contrary to what is usually observed at Kilauea⁶.

The difference in the seismicity of the two volcanoes may be related to the more complex structure of Kilauea which has several centres of activity. The internal structure of Piton de la Fournaise is not well documented. A detailed study using magnetic data⁷ suggests the existence of a small kilometre-sized magma chamber ~1 km beneath the summit area. Shallow seismicity accompanied by inflation is presumably related to the extension of the central dyke complex located immediately below the summit⁷.

Eruptive activity occurred in three phases roughly 1 month long (Table 1). Each phase conformed to the same pattern: volcanic tremors for 3–4 h at the start, and virtually no seismicity during the rest of the phase when the greatest volume of lava is erupted. This indicates that tremor activity is due to the initiation of flow.

Tholeiitic lavas came out of fissures which opened along well known NE–SW and NW–SE axes of weakness. They were essentially aphyric⁸, contrary to the previous eruptions of April 1977 and July 1979. The three eruptive phases were similar, and characterized by intense degassing with lava fountains reaching heights of 70–100 m. After a couple of days, the activity became limited to a few fissures where spatter cones grew. One or two days later, only one fissure was productive and a volcanic cone was built there. The velocity of lava flow was estimated visually to be close to 4 m s⁻¹ at the vent.

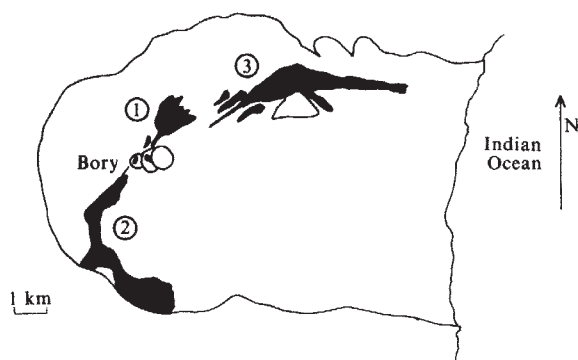


Fig. 4 Map of the lava flows produced by the three eruptive phases.

Lavas produced during this eruption cover large areas (Table 1, Fig. 4) and the total volume is $\sim 10 \times 10^6 \text{ m}^3$, using a mean flow thickness of 5 m. Lava flows were essentially of the aa, grats and echaudé types which are typical of the island, with some pahoe-hoe late in each eruptive phase.

Several phreatomagmatic explosions were observed during the first phase, and many during the third, in relation with heavy rainfall.

This is IGP contribution no. 589.

Received 18 January; accepted 5 April 1982.

1. Wadge, G. *Nature* **288**, 253–255 (1980).
2. Bachelery, P. thesis, Univ. Clermont Ferrand (1981).
3. Blum, P. A., Gaulon, R., Lalanne, F. X. & Ruegg, J. C. *C.R. hebdomadaire. Séances. Acad. Sci., Paris* **292**, 1449–1455 (1981).
4. Koyanagi, R. Y., Meagher, K., Klein, F. W. & Puniwai, G. S. *Hawaiian Volc. Obs. Summary* (1976).
5. Nakata, J. S., Tanigawa, W. R. & Klein, F. W. *Hawaiian Volc. Obs. Summary* (1979).
6. Koyanagi, R. Y., Endo, E. T. & Ward, P. L. *Geophys. Monogr.* **19**, 169–173 (1976).
7. Lenat, J. F. & Aubert, M. *J. Volcan. geotherm. Res.* (in the press).
8. Chevallier, L., Lalanne, F. X., Bachelery, P. & Vincent, P. M. *C.R. hebdomadaire. Séances. Acad. Sci. Paris* **293**, 187–190 (1981).

Prehistoric soil and vegetation development on Bodmin Moor, southwestern England

E. Maltby & C. J. Caseldine

Geography Department, University of Exeter, Exeter, Devon EX4 4RJ, UK

Archaeological excavation of barrows in the St Neot Valley on the southern flanks of Bodmin Moor revealed buried soils developed in the gravelly granitic head typical of the upland block. By comparing these soil profiles with adjacent soils we have obtained evidence for and now describe dramatic prehistoric changes in upland soils and vegetation attributed previously to various combinations of human influences and climatic events. Direct pedological and palaeoecological evidence supports the theory that brown soils existed in pre-Bronze Age times in southwestern England at a site subsequently developing an iron pan (Bf) and accumulating a strongly acid peaty surface.

The barrows CR11 and CR14 revealed by the excavation have yielded statistically inseparable radiocarbon dates of 1550–1650 and 1560–1630 b.c. (ref. 1) respectively and occupy opposing valley shoulder positions <500 m apart (Fig. 1). The CR14 complex stands in acid grass–heath moorland characteristic of the uplands of south-west England and in which *Molinia caerulea* and *Calluna vulgaris* are locally prominent. The

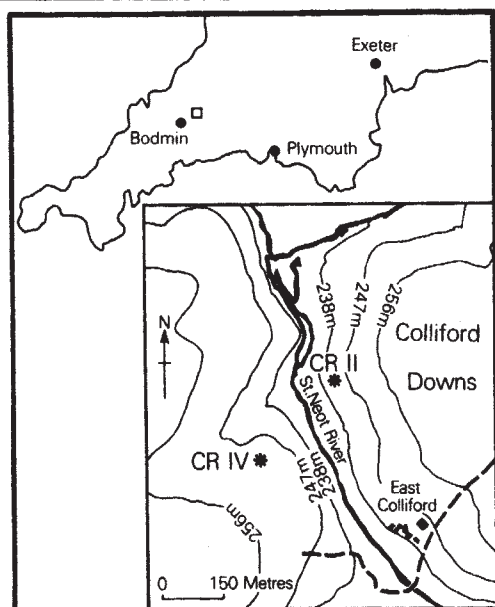


Fig. 1 Location of Bronze Age barrows in the St Neot Valley, Bodmin Moor.

present day soil is a distinctive iron pan stagnopodzol comprising 15–20 cm peaty surface, a variable but stony Ah, grey mottled or pallid E(g) above a thin but wavy and indurated iron pan (Bf) bounding ochreous then brown B_s horizons which merge at 50–60 cm into paler B/C material. CR II occupies a slope enclosed and converted to improved pasture at least since the seventeenth century. Associated soils have a well-mineralized organic surface, commonly lack an E horizon (which has been lost probably by cultivation), but still possess an iron pan (occasionally discontinuous) and strongly developed B_s horizon. It is a cultivated variant of the iron pan stagnopodzol which has been variously mapped and described elsewhere as the Hexworthy Series^{2,3}.

The profiles beneath two barrow structures of the CR IV complex (CR IVa and CR IVc) are identical to the present day soil except that the original surface peat is attenuated at the

old land surface. This is explained by the incorporation of peaty turves into the structure of the barrows. However, a strongly contrasting relationship emerges from the sub-CR II soil. The mineralogical profile comprises three morphologically distinct horizons above B/C and C(x) material (Fig. 2a). Unlike the profile developed within and at the margins of the barrow, under CR IV and associated with the present day surface, there is no organic surface, bleached E_a or E_g horizon or morphological evidence for an iron pan. It seems possible that an original turf layer was removed in the course of barrow construction but otherwise the profile is intact. Horizon notation for the buried profile follows Avery⁴ but non-pedogenic nomenclature is preferred for our discussion in view of the potential difficulties in interpretation of buried horizons.

Relatively high values for sodium dithionite extractable iron (Fe_{res}) which are remarkably consistent with depth suggests a strong affinity⁵, if not genetic link, between the buried soil and a brown earth type progenitor (Fig. 2b). Such crystalline 'aged' Fe only remains in the absence of active biochemical agencies of remineralization usually associated with acid organic surfaces⁵ and/or leaf leachates^{6,7}. Thus alternatively the pattern of Fe_{res} may be an artefact of burial and protection from fresh sources of complexing or chelating radicals. However, alkaline potassium pyrophosphate extracts (Fe_{ppt}, Fe_{sol}, Al) show particular enrichment of 'active' amorphous Fe in horizon II and confirms general podzolization of the profile. Iron which can be precipitated by NH₃ (Fe_{ppt}) is the dominant fraction in all but the lowest levels sampled. That fraction of the extract remaining soluble (Fe_{sol}) is described by Bascomb as the most active form of Fe linked in 'fresh' organic complexes⁵. Amounts are generally low. This might be expected if the profile were protected substantially from contemporary leaching under the large granite blocks of the barrow but is also a typical feature of brown soils. The proportion of 'active' Fe (Fe_{ppt} + Fe_{sol} = Fe_{ext}) to total Fe (Fe_{ext} + Fe_{res}) changes with depth in a pattern confirming active enrichment of II and a change to mainly 'aged' iron in underlying B/C material.

A marked peak in active Fe is associated with organic rich infill of a distinct system of channels. The channels are unusual for several reasons: they contain black peaty fill including bleached quartz grains similar to the modern moorland and

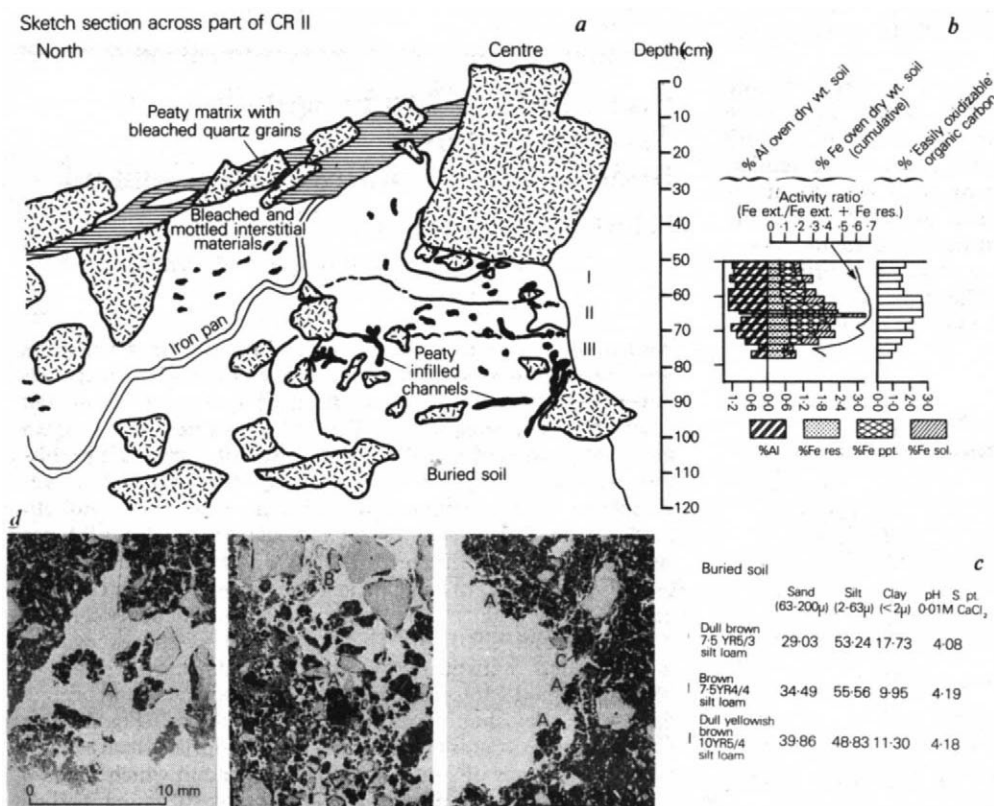


Fig. 2 a, Section across part of CR II revealing buried soil especially well preserved beneath large granite block. Note iron pan soil well developed in barrow structure but Bf horizon not extending into buried profile. b, Distribution of pyrophosphate extractable Fe, Al and dithionite extractable Fe and 'easily oxidizable' organic carbon (rapid dichromate digestion) in buried profile. c, Colour (Munsell notation), textural characteristics (wet sieving and sedigraph analysis), and pH ('sticky point' method on fresh material) of buried horizons. d, Thin sections across channel features. A, probable earthworm cast material (highly organic); B, root epidermis enclosing small animal droppings; C, void with fine organic detrital material.

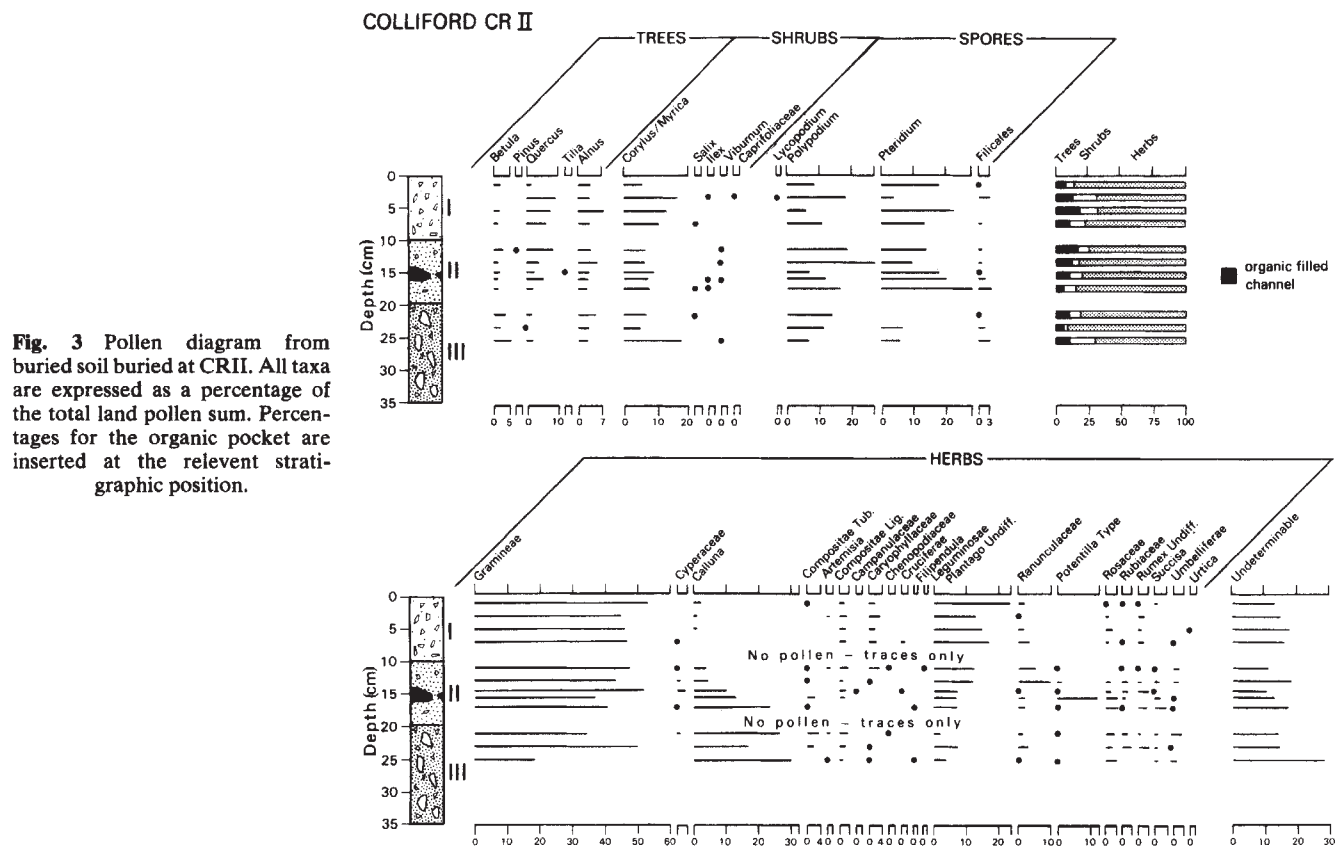


Fig. 3 Pollen diagram from buried soil buried at CRII. All taxa are expressed as a percentage of the total land pollen sum. Percentages for the organic pocket are inserted at the relevant stratigraphic position.

Table 1 Dominant and characteristic pollen taxa from the soil horizons of CRII

Horizon	Pollen taxa
I	Dominated by Gramineae and <i>Plantago</i> with <i>Quercus</i> , <i>Betula</i> and <i>Corylus/Myrica</i> comprising 30% T.L.P. Variable but high frequencies for <i>Polypodium</i> and <i>Pteridium</i> .
II	Dominated also by Gramineae and <i>Plantago</i> at reduced levels with reduced <i>Quercus</i> and <i>Corylus/Myrica</i> . Increasing <i>Calluna</i> and greater species diversity including Ranunculaceae, <i>Potentilla</i> type, Rosaceae, Rubiaceae, <i>Succisa</i> and <i>Rumex</i> . Highest values for <i>Polypodium</i> and <i>Pteridium</i> .
III	Dominated by lower values for Gramineae, and <i>Calluna</i> . <i>Plantago</i> , <i>Pteridium</i> , <i>Polypodium</i> and all A. P. are reduced. At the lowest sample there is a peak for <i>Corylus/Myrica</i> .
Organic pocket	Similar assemblage to surrounding horizon with dominance of Gramineae, <i>Plantago</i> and <i>Calluna</i> with <i>Corylus/Myrica</i> , <i>Alnus</i> and <i>Quercus</i> .

barrow surface; they exhibit remarkably little variation in internal diameter (mean = 9.84 mm, $s = \pm 1.12$ mm, $v = 12\%$, $n = 25$) except for occasional enlarged pockets; where channels split or bifurcate they often do so in an upward direction; the internal walls are commonly smoothed; they lack smaller lateral systems and macro-remains of recognizable plant debris such as bracken rhizomes. Thin sections reveal the presence of possible earthworm casts within and lining the channels together with evidence in pores of substantial small animal activity (Fig. 2d). We propose that these channels are 'fossilized' burrows of lumbricid earthworms with which they compare in permanence, size and depth⁸. Absence of a highly acidic peat or mor humus surface is necessarily implied by their establishment. The population persisted until soil conditions degraded to a point which allowed accumulation of an acid peaty surface as indicated by the character of the fill in the abandoned burrows. Earthworm activity could be responsible for the much finer

texture of horizon I (A/Ah?) and also the general homogeneity of the pollen record within this horizon can be explained by active faunal mixing (Fig. 2c). Pollen in the channels themselves is better preserved in the more acid and recent fill than generally in the buried profile. An assemblage with up to 15% *Calluna* confirms the idea of infilling occurring when a peaty surface had begun to develop in association with acidophilous heathland communities. Large earthworms died out at least a short time after 1560 b.c. because several channels extend into the barrow structure. Note in this context that at the experimental earthwork on Overton Down, constructed in 1960, earthworm activity rapidly extended well up into the mound from the buried soil surface. By 1964 worm casts are recorded in interstitial spaces throughout the experimental bank⁹.

A pollen diagram has been constructed for the buried soil despite problems caused by generally poor pollen preservation (Fig. 3). Changes in the gross character of pollen assemblages between the three horizons are summarized in Table 1. Pollen evidence suggests that the immediate pre-barrow soil developed in a relatively open grassland environment. Increasing *Calluna* down the profile, however, indicates that heather was an important component of the local vegetation before grassland development. It may have comprised the under-storey or gaps in scrubby woodland dominated by hazel or oak already substantially modified by Neolithic human activity. Indiscriminate burning and/or excessive grazing, together with more cultivation could have led to its demise as pressure for land increased into the Bronze Age.

It has been argued that the Hexworthy Series on Dartmoor has developed from an earlier brown earth type under oak forest¹⁰. Enhanced acidification and leaching have been associated with Neolithic clearance on land generally above 300 m but because of the temporary nature of such clearings and the inferred frequent regeneration of woodland it is by no means universally agreed that there was any long lasting effect on soil development. Pre-Bronze Age podzolization and iron pan development at ~250–300 m elevation is inferred for Dartmoor and certainly pre-dates barrow construction of this period, including that of CRIV elsewhere on the granite uplands of

southwestern England^{13,14}. The present study confirms for the area immediately adjacent to CR11, however, that soil changes leading to iron pan formation and accumulation of a peaty surface post-dates this period. Nevertheless, podzolization was already well enough advanced to produce a brown (possibly humic) podzol profile. A wide variety of buried soils ranging from 'unleached' profiles to thin iron pan podzols were described in association with Bronze Age barrows elsewhere in upland Britain and differences explained in terms of very rapid soil changes occurring within this period of prehistoric activity¹⁵. High rates of differential pedogenesis cannot be dismissed entirely on Bodmin Moor. However, soil changes are not necessarily time-synchronous even over short distances and the effects of slope aspect, subtle differences in elevation, gradient and hydrology together with any differential influences of prehistoric land management should not be ignored as causes of significant time delays in the processes of conversion of brown soils to iron pan stagnopodzols.

Frances Griffith has coordinated the archaeological programme of which this study has been part. We thank Nigel Fenn and Anne Alderton, financed through the MSC Job Creation Programme, for research assistance and Dr Peter Bullock for preparing thin sections and for his helpful interpretation. South West Water are thanked for site access. Figures are by the drawing office and photographic unit, Geography Department, Exeter University.

Received 1 February; accepted 7 April 1982.

1. Griffith, F. *Corn. Archaeol.* **18**, 56 (1979).
2. Clayden, B. *Soils of the Exeter District* (Soil Survey England and Wales, Harpenden, 1971).
3. Staines, S. J. *Soil in Cornwall I* (Soil Survey Record No. 34, Harpenden, 1976).
4. Avery, B. W. *Soil Classification for England and Wales* (Higher Categories) (Soil Survey Technical Monograph 14, Harpenden, 1980).
5. Bascomb, C. L. *J. Soil Sci.* **19**, 251–268 (1968).
6. Bloomfield, C. *Rep. Welsh Soils Discuss. Group* **11**, 112–123 (1970).
7. Davies, R. I. *Rep. Welsh Soils Discuss. Group* **11**, 133–142 (1970).
8. Edwards, C. A. & Lofty, J. R. *Biology of Earthworms* (Chapman & Hall, London, 1972).
9. Jewell, P. A. & Dumbleby, G. W. *Proc. prehist. Soc.* **32**, 313–342 (1966).
10. Staines, S. J. thesis, Univ. Bristol (1973).
11. Simons, I. G. *Proc. prehist. Soc.* **35**, 203–219 (1969).
12. Staines, S. J. *Proceedings Devon archaeol. Soc.* **37**, 21–47 (1979).
13. Staines, S. J. *Cornish Archaeol.* **14**, 66–67 (1975).
14. Brisbane, M. & Clews, S. *Cornish Archaeol.* **18**, 33–56 (1979).
15. Dumbleby, G. W. *The Development of British Heathlands and Their Soils* (Oxford Forestry Mem. No. 23, 1962).

SO₂ pollution reduces the freezing resistance of ryegrass

Alan W. Davison & Ian F. Bailey

Department of Plant Biology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK

Several authors^{1–3} have observed that the effects of SO₂ on growth are more severe in winter and Davies⁴ has shown that this may be due to low irradiance and the short photoperiod reducing the plants' capacity to detoxify the pollutant. However, in winter plants are also subjected to potentially lethal stresses such as freezing. The important forage grass, *Lolium perenne*, shows only poor resistance to winter temperatures, hence, any effect of pollution on hardiness could be of considerable significance. Here we demonstrate that SO₂ reduces resistance to sub-zero temperatures and that this effect is influenced by mineral nutrition. A reduction in freezing tolerance may contribute to rapid selection of SO₂-resistant genotypes, particularly in newly sown leys.

Three experiments were performed to determine the effects of SO₂ on freezing resistance. In the first, six 1-week-old seedlings of *L. perenne* cv. S23 were planted in each of 90 7-cm plastic pots containing standard garden potting compost and these were fumigated for 3 weeks in wind tunnels. During this phase of the experiment the day length was 18 h, irradiance 32 W m⁻² and temperature 15 °C. Controls received <15 µg SO₂ per m³ and treated plants 250 µg per m³. SO₂ was monitored using a Meloy SA 285 sulphur analyser.

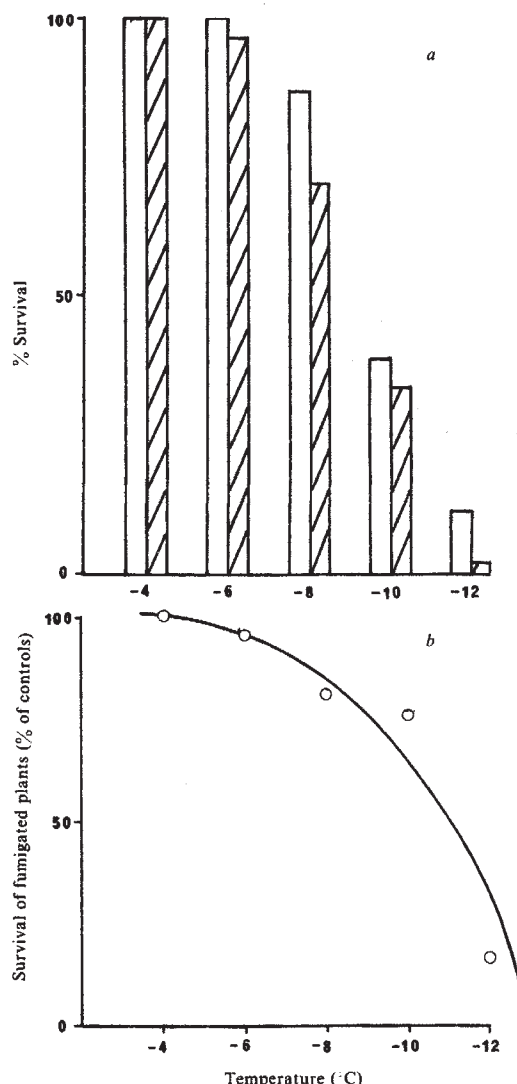


Fig. 1 Survival of *L. perenne* exposed to sub-zero temperatures and SO₂. a, Percentage survival of control (open bars, <15 µg SO₂ m⁻³) and fumigated (cross-hatched bars, 250 µg SO₂ m⁻³) plants. LT₅₀ of controls = -9.7 ± 0.4 °C, that of fumigated plants = -8.8 ± 0.3 °C. b, Survival of fumigated plants expressed as a percentage of controls.

After this initial period, fumigation conditions were kept the same except that the temperature was reduced to 2 °C and the daylength to 10 h to 'harden' the plants. After 2 weeks of such treatment, all the plants were removed from the wind tunnels and placed in randomized blocks in a modified deep freeze (daylength, 10 h and irradiance, 55 W m⁻²). They were not fumigated during this phase. Using a cam-operated temperature control, the plants were held at 0 °C during the first day, followed by exposure during the night to a fall in temperature of 2 °C h⁻¹ to a minimum of -4 °C. This minimum was kept constant for 4 h then the temperature was allowed to rise at 2 °C h⁻¹ back to the daytime temperature of 0 °C. On subsequent nights the same regime was followed except that the minimum was lowered to give -4, -6, -8, -10 and -12 °C on successive nights. Nine replicate fumigated and control pots were removed at random each day and kept without fumigation at 2 °C in the wind tunnels until all the plants had been removed from the freezer. Then the plants were returned to the same environmental conditions as those during the 3-week pre-hardening period. After 5 weeks of recovery, the plants were scored for survival (post-freezing production of new root, leaf or tiller). The results are shown in Fig. 1.

Survival of the fumigated plants was consistently poorer than controls and probit analysis showed that the LT₅₀ values (tem-

Table 1 Influence of soil sulphur and nitrogen on the effects of SO₂ fumigation in *Lolium perenne*

	Low S/ low N	Low S/ high N	High S/ low N	High S/ high N
Control ($<15 \mu\text{g SO}_2 \text{ m}^{-3}$)	239* (22)	424† (18)	235* (22)	487‡ (16)
+SO ₂ ($250 \mu\text{g SO}_2 \text{ m}^{-3}$)	208* (18)	400† (20)	202* (12)	400† (17)

Mean dry wt (mg) of shoots of *L. perenne* cv. S23. Means identified by different symbols are significantly different ($P < 0.02$) from each other. Standard errors are shown in parentheses.

peratures at which 50% of the plants were killed) were respectively -8.8 ± 0.3 and -9.7 ± 0.4 °C. Figure 1b shows that when survival of fumigated plants was plotted as a percentage of controls the difference between the treatments increased with decreasing temperature.

As the supply of nitrogen and other elements influences⁵⁻⁹ the effects of SO₂, we performed experiments to determine the effects of soil nitrogen and sulphur on the SO₂-freezing interaction. The first of these experiments was a fumigation-growth study which was used to select contrasting nutrient treatments for use in the second experiment. In the former, newly germinated seedlings of *L. perenne* cv. S23 were sown at a density of 5 per 7-cm plastic pot containing washed, coarse sand. All treatments received deionized water each day and 10 ml of nutrient solution per pot twice a week. Four nutrient treatments were used (Table 1) to give low soil S-low N, low S-high N, high S-low N and high S-high N. All nutrient solutions contained (mg l^{-1}): KH₂PO₄ 180, CaCl₂·2H₂O 367, MgCl₂·6H₂O 313 and trace elements, together with different concentrations of Na₂SO₄ and NH₄NO₃. The low and high sulphur solutions contained 21.3 and 213 mg l^{-1} Na₂SO₄, giving 4.8 and 48 mg S l^{-1} , respectively, while the low and high nitrogen solutions contained 40 and 400 mg l^{-1} NH₄NO₃, giving 14 and 140 mg N l^{-1} , respectively. Twelve replicate pots were used for

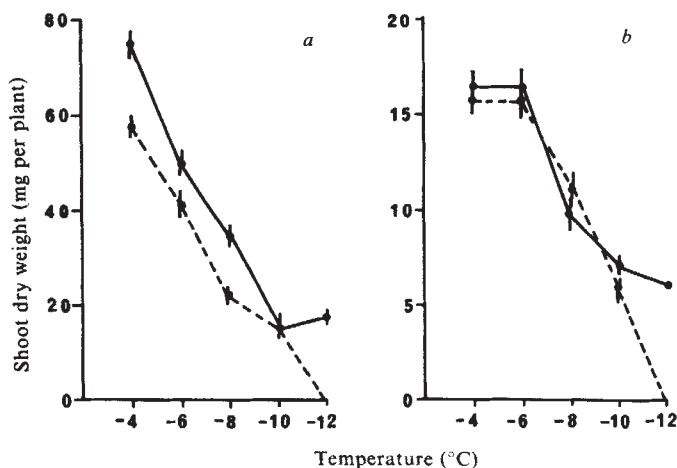


Fig. 3 Dry weight (mg per plant) of *L. perenne* shoots exposed to sub-zero temperatures and SO₂. a, Control (—) and fumigated (---) plants on high sulphur-high nitrogen medium. b, Plants on low sulphur-low nitrogen medium. Vertical bars represent standard errors.

each treatment. Immediately after sowing, the pots were placed in randomized blocks in wind tunnels and exposed to the same environmental conditions as in the pre-hardening period of the first experiment. After 5 weeks of fumigation, the plants were harvested, dried at 105 °C and weighed. The results are shown in Table 1.

In both control and fumigated treatments, an increase in nitrogen concentration significantly increased yield but the effects of soil sulphur were more complex. Increasing the sulphur concentration increased yield only in the presence of high nitrogen concentration in clean air—that is, sulphur was limiting only when the nitrogen concentration was high and SO₂ was low. SO₂ did not appear to compensate for low soil sulphur. With high substrate sulphur and nitrogen, SO₂ caused a significant ($P < 0.02$) reduction (18%) in shoot weight. Based on these results, the two extreme nutrient treatments (low S-low N and high S-high N) were used in the final experiment, which examined the interaction of SO₂, substrate sulphur and nitrogen, and freezing injury. Seedlings were grown in sand culture in the same environmental conditions and regime as in the previous experiment. After 5 weeks of recovery from freezing, the plants were scored for survival, harvested, dried and weighed (Figs 2, 3).

The reaction of plants given high substrate sulphur and nitrogen to sub-zero temperatures (Fig. 2a) confirmed the results of the first experiment in that SO₂ reduced both survival and the LT₅₀ (from -8.4 ± 0.4 °C to -7.6 ± 0.2 °C). However, the difference between control and fumigated plants was greatly reduced (Fig. 2b) when they were grown on the low sulphur-low nitrogen medium. The poor winter hardiness of *L. perenne* was reflected not only in the LT₅₀ but also in the dry weights. In both nutrient regimes and for both control and fumigated plants, lowering of the temperature caused a progressive decrease in shoot weight but the weight of fumigated plants on the high sulphur-high nitrogen medium was consistently less than controls. On the low sulphur-low nitrogen medium, SO₂ had no effect on shoot weight except at the lowest temperature, where no fumigated plants survived. Root weights showed the same pattern as the shoots (data not shown).

The possibility that pollutants might affect winter hardiness was first mentioned by Bleasdale¹⁰ in 1952 but until now there has been no direct evidence that SO₂ affects any of the attributes that constitute what we call hardiness. A reduction in survival at temperatures that are commonly recorded in British winters has several implications, particularly in relation to efforts to quantify the economic impact of SO₂. Several authors¹¹⁻¹³ have commented on the difficulties involved in extrapolating results obtained in the laboratory to the field situation because of the

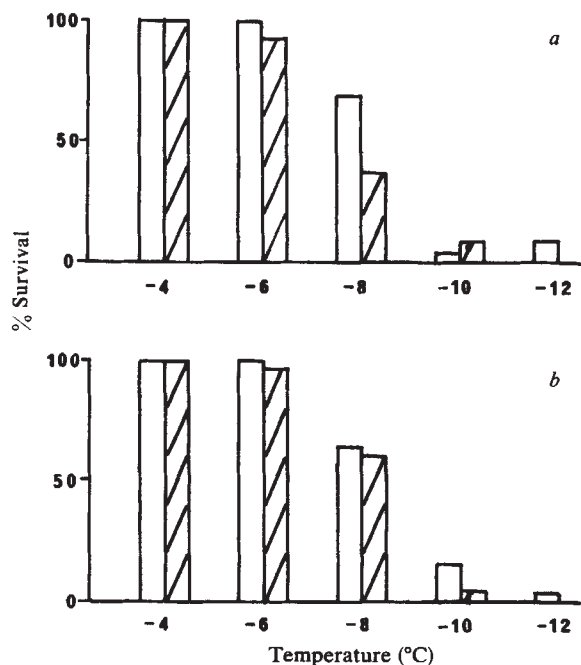


Fig. 2 Survival of *L. perenne* exposed to sub-zero temperatures and SO₂. a, Percentage survival of control (open bars, $<15 \mu\text{g SO}_2 \text{ m}^{-3}$) and fumigated (cross-hatched bars, $250 \mu\text{g SO}_2 \text{ m}^{-3}$) plants on high sulphur-high nitrogen medium. LT₅₀ of controls = -8.4 ± 0.4 °C, that of fumigated plants = -7.6 ± 0.2 °C. b, As for a, but plants were given low sulphur-low nitrogen medium. LT₅₀ of controls = -8.5 ± 0.4 °C, compared with -8.4 ± 0.3 °C for fumigated plants.

effects of variable environmental factors such as light, genetic variation and interactive effects of other pollutants. The finding of an effect of SO₂ on freezing resistance and its modification by mineral nutrition adds to the difficulties and reinforces the view that in order to quantify effects as they occur in the field, experiments must involve exposure of plants at all times of the year and to all the normal range of environmental conditions.

There is good evidence¹⁴⁻¹⁸ that in areas subjected to SO₂ pollution there exists natural selection of tolerant populations. The rate of selection has not been fully established¹⁸ but if SO₂ reduces freezing resistance then this could promote extremely rapid selection, especially in re-seeded leys. It is possible that selection might be so rapid that the effects of SO₂ on the production of re-seeded grasslands might be transient and limited to a short period immediately after sowing.

We thank Alcan (Lynemouth) for financial assistance and Dr W. S. Stewart for the loan of equipment.

Received 25 January; accepted 2 April 1982.

1. Bell, J. N. B. & Clough, W. S. *Nature* **241**, 47-48 (1973).
2. Ashenden, T. W. & Mansfield, T. A. *Nature* **273**, 142-143 (1978).
3. Young, J. E. & Matthews, P. *Envir. Pollut. B* **2**, 353-365 (1981).
4. Davies, T. *Nature* **284**, 483-485 (1980).
5. Cowling, D. W. & Jones, L. H. P. *Soil Sci.* **110**, 346-354 (1970).
6. Leone, I. A. & Brennan, E. J. *Air Pollut. Control Ass.* **22**, 544-547 (1972).
7. Cowling, D. W., Jones, L. H. P. & Lockyer, D. R. *Nature* **243**, 479-480 (1973).
8. Cowling, D. W. & Lockyer, D. R. *J. exp. Bot.* **29**, 257-265 (1978).
9. Ayazloo, M., Bell, J. N. B. & Garsed, S. G. *Envir. Pollut. A* **22**, 295-307 (1980).
10. Bleasdale, J. K. A. *Nature* **169**, 376-377 (1952).
11. Bell, J. N. B. *Nature* **284**, 399-400 (1980).
12. Mansfield, T. A. & Freer-Smith, P. H. *Biol. Rev.* **56**, 343-368 (1981).
13. Davison, A. W. *Br. Ecol. Soc. Symp.*, Essex 1980 (in the press).
14. Taylor, G. E. & Murdy, W. H. *Bot. Gaz.* **136**, 212-215 (1975).
15. Bell, J. N. B. & Mudd, C. H. in *Effects of Air Pollutants on Plants* (ed. Mansfield, T. A.) 87-103 (Cambridge University Press, Cambridge, 1976).
16. Wellburn, A. R., Capron, T. M., Chan, H.-S. & Horsman, D. C. in *Effects of Air Pollutants on Plants* (ed. Mansfield, T. A.) 105-114 (Cambridge University Press, Cambridge, 1976).
17. Horsman, D. A., Roberts, T. M. & Bradshaw, A. D. *Nature* **276**, 493-494 (1978).
18. Bell, J. N. B., Ayazloo, M. & Wilson, G. B. in *Proc. 2nd Eur. Ecological Symp.*, Berlin (in the press).

'Wasted', a new mutant of the mouse with abnormalities characteristic of ataxia telangiectasia

L. D. Shultz, H. O. Sweet, M. T. Davisson & D. R. Coman

The Jackson Laboratory, Bar Harbour, Maine 04609, USA

Research into the basic mechanisms underlying the development of complex multi-system diseases has been facilitated by the study of single gene mutations in experimental animals. Ataxia telangiectasia is such a multifaceted genetically determined disease for which no animal model has been described. It occurs in children and is characterized by varying degrees of immunodeficiency, cerebellar ataxia, oculocutaneous telangiectasia, chronic sinopulmonary disease, endocrine abnormalities, chromosomal aberrations, and a high incidence of neoplasms¹⁻³. Although the clinical manifestations have been well documented, the cause of the disease is unknown. The pleiotropic nature of the ataxia telangiectasia gene is very complex. Many aspects of this disease, including impairment of tissue differentiation in embryological development are not amenable to direct study, thus an animal model for ataxia telangiectasia would be valuable in elucidating the underlying mechanisms, and in developing effective preventive or therapeutic measures. We have recently found a spontaneous mutation of the mouse called 'wasted' (*wst*) that shows pathological changes similar to those of ataxia telangiectasia. This mutation can be recognized at 3 weeks of age by neurological abnormalities. We report here that affected animals show pathological changes in the central nervous and lymphoid systems and exhibit a high degree of spontaneous and γ ray-induced chromosomal damage.

The autosomal recessive mutation *wst* occurred spontaneously in 1972 in the inbred HRS/J colony maintained in

the Mouse Mutants Stocks Center at the Jackson Laboratory⁴. Homozygotes (*wst/wst*) can be recognized at 20 days of age by neurological abnormalities, including tremor and uncoordinated body movements. Affected mice develop progressive paralysis and do not survive beyond 30 days of age. As wasted mice are short-lived, the mutation was crossed to a segregating background (C3HeB/FeJ $\times$ C57BL/6J) to increase viability through 'hybrid vigor'. We then studied the effect of the *wasted* mutation on the segregating hybrid background. We are presently backcrossing the *wasted* mutation to the C57BL/6J strain. Although *wst* has not yet been mapped, it does segregate independently of markers on chromosomes 1-11, 13-15, 17 and 19; further linkage tests are in progress. Allelism tests using *ax¹*, *bc*, *cri*, *dw*, *dy*, *mdf*, *med¹*, *myd*, *nr*, *pcd*, *tn*, *twi* and *wl* were all negative. (For a description of these genes, see ref. 5.)

While mice normally show rapid growth and weight gain between 20 and 30 days of age, *wasted* homozygotes show no weight increase after 20 days. Of particular interest was the finding that, by 28 days of age, *wasted* mice have a marked lymphoid hypoplasia with significantly decreased ($P < 0.05$, Student's *t*-test) spleen, thymus, and lymph node to body weight ratios (Table 1). Analysis of peripheral blood in 17 pairs of mutant and control mice revealed that 3-4-week-old *wasted* homozygotes have a threefold decrease in numbers of circulating leukocytes ($2.8 \pm 0.52 \times 10^6 \text{ ml}^{-1}$ for *wst/wst* mice as opposed to $8.7 \pm 1.06 \times 10^6 \text{ ml}^{-1}$ for $\pm$ littermates). To determine whether the *wst* mutation affects development of both Thy-1-bearing cells (T cells) and surface immunoglobulin-bearing cells (B cells), spleen cell suspensions were analysed by indirect immunofluorescence microscopy⁶ using monoclonal anti-Thy 1.2 antibody (cell culture supernatant from hybridoma HO-13-4.9)⁷ and fluorescein-conjugated goat anti-mouse immunoglobulin. Preliminary analysis of nine pairs of *wasted* and littermate control mice (20-28 days old) showed no significant effect of the *wst* mutation on percentages of T and B cells (data not shown). Thus, the lymphoid hypoplasia is not restricted to a single major lymphocyte subpopulation. Similarly, in ataxia telangiectasia patients, percentages of circulating B cells are normal while those of T cells are within the lower normal limits⁸.

As patients having ataxia telangiectasia show impairment of delayed-type hypersensitivity (DTH), we examined the DTH response of *wasted* mice to sheep red blood cells (SRBC). Homozygotes (*wst/wst*) had significantly decreased DTH responses 4 days after priming compared with $\pm$ littermate controls, as measured by the specific increase in footpad thickness 24 h after injection of an eliciting dose (10^8) of SRBC⁹ (data not shown).

Histological examination of tissues from 10 pairs of *wasted* and littermate control mice (22-29 days old) revealed abnormalities in the central nervous and lymphoid systems of mutant mice. The brain showed degeneration of Purkinje cells and focal demyelination in the cerebellar cortex (Fig. 1); focal demyelination was also found in the ventral columns of the spinal cord of *wasted* mice. Purkinje cell degeneration and focal demyelination have also been described in association with ataxia telangiectasia^{1,3,10}. Marked hypoplasia was found in both the thymus-dependent and thymus-independent areas of the lymphoid system. The follicles of the lymph node and spleen

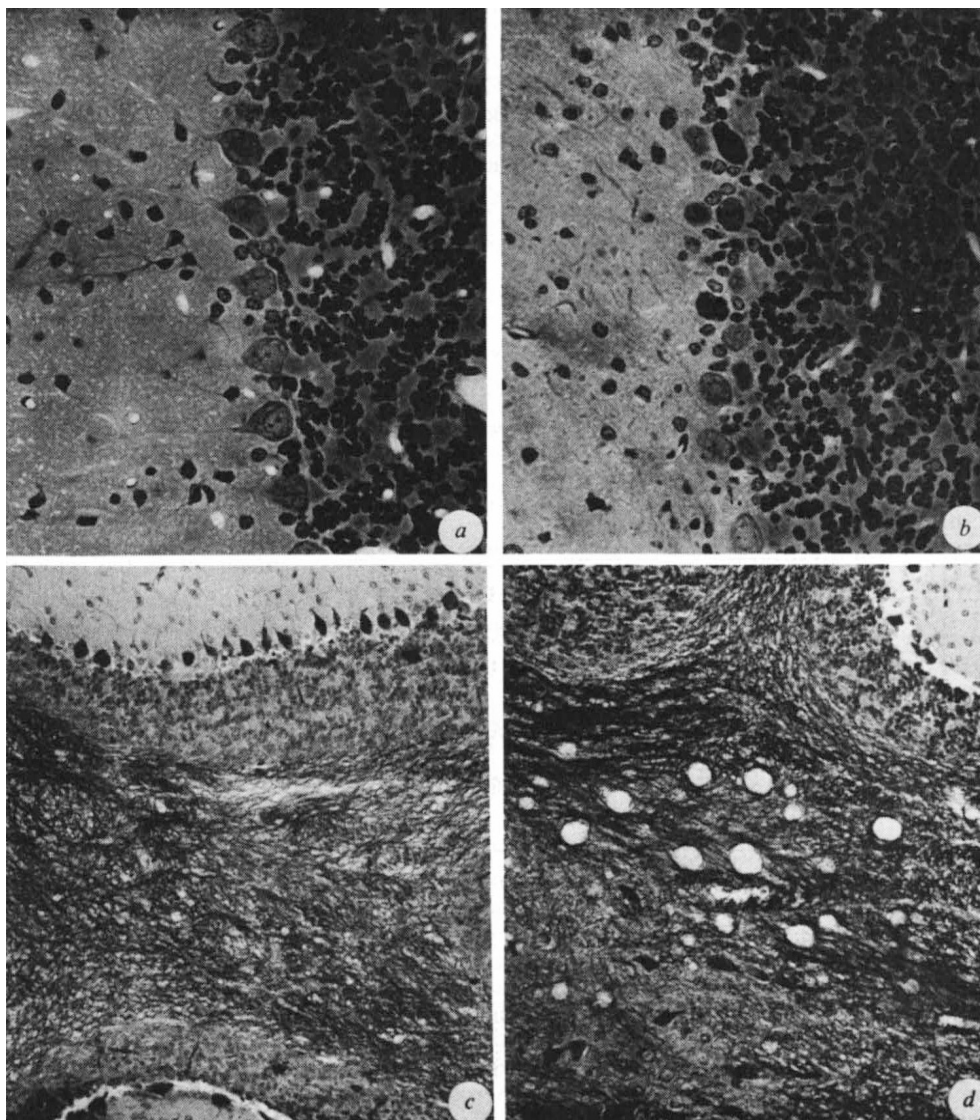
Table 1 Body weight and lymphoid organ/body weight ratios of *wasted* mice

Genotype	Body weight (g)	Lymphoid organ weight (mg)/body weight (g)		
		Thymus	Spleen	Lymph node*
<i>wst/wst</i>	7.8 $\pm$ 0.51	2.4 $\pm$ 0.34	2.0 $\pm$ 0.23	0.19 $\pm$ 0.03
$\pm$	14.8 $\pm$ 0.61	5.2 $\pm$ 0.29	4.3 $\pm$ 0.17	0.27 $\pm$ 0.02

Values are mean $\pm$ s.e.m. ($n = 8$). Lymphoid organs were taken from 28-day-old *wasted* and littermate control mice of the same sex.

* Right brachial lymph nodes were weighed.

Fig. 1 Sections of brain from 29-day-old wasted and littermate control mice. *a*, Cortex of cerebellum from littermate control showing normal Purkinje cells; arborescent fibrils are evident. *b*, Cortex of cerebellum from *wst/wst* mouse; Purkinje cells appear shrunken and deficient in fibrils (luxol fast blue-cresyl violet, $\times 140$). *c*, Cerebellum from littermate control showing normal myelination. *d*, Cerebellum from *wst/wst* mouse showing advanced myelin loss (luxol fast blue-cresyl violet, $\times 70$).



were poorly developed, Peyer's patches were few and small, and there was a reduction in cellularity of the thymic cortex (Fig. 2). Thymic hypoplasia and lymphocyte depletion in the cortical and paracortical areas of lymph nodes are common features of ataxia telangiectasia³. No gross or histological evidence of telangiectases was found in wasted mice, but development of such vascular abnormalities is thought to require exposure to sunlight¹¹.

As ataxia telangiectasia is characterized by spontaneous and induced chromosomal aberrations, such changes were sought in wasted mice. We examined chromosomes in bone marrow cells collected from nine pairs of *wasted* and littermate control mice (23–28 days old). Cells were not G-banded so that they could be scored for chromosome gaps. At least 50 cells from each mouse were scored blindly for chromosome damage. The *wst/wst* mice showed a fourfold greater incidence of chromosomal damage than their littermate controls ($8.8 \pm 1.69\%$ of cells from *wst/wst* mice compared with $2.2 \pm 0.77\%$ of cells from $\pm/\pm$ littermates showed chromosomal damage). The damage included chromosome and chromatid breaks, gaps and fragments, implying either increased chromosome breakage or failure of breaks to rejoin. Preliminary data suggest that wasted mice may also be more susceptible to radiation-induced chromosomal injury. Two *wst/wst* mice and two littermate controls were irradiated with 200R from a Shepard Mark 1 irradiator (¹³⁷Cs). After 18 h the mice were treated with colchicine, and chromosomes were prepared from bone marrow cell suspensions; 80% of 30 metaphase cells in *wst/wst* mice

showed damage of some sort while the controls revealed only minor damage in 30% of the cells scored.

Thus, homozygosity for the wasted mutation results in impaired neurological and immunological development as well as increased levels of spontaneous and γ ray-induced chromosomal damage. Whether the chromosomal aberrations in *wasted* mice are associated with defective DNA repair, as has been shown in cell lines isolated from individuals with ataxia telangiectasia¹, has not been determined. Elucidation of the

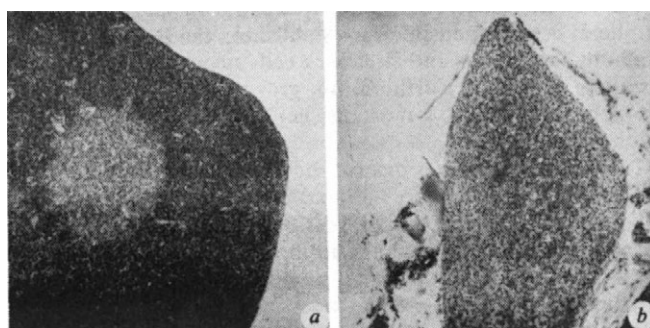


Fig. 2 Thymuses from 29-day-old wasted and littermate control mice. *a*, Littermate control showing normal thymic structure; *b*, 'wasted' thymus showing depletion of cells from cortex (haematoxylin-eosin, $\times 25$).

mechanism of action of the wasted mutation may further our understanding of ataxia telangiectasia.

We thank Cheryl Bailey for technical assistance. This work was supported by NIH research grants CA20408 and GM27102, NSF grant DEB79-26708, and ACS grant CD-34U. *Note added in proof:* Associated with the focal demyelination in the brain and spinal cord of wasted mice are numerous dilated capillaries.

Received 15 March; accepted 13 April 1982.

1. Paterson, M. C. & Smith, P. J. A. *Rev. Genet.* **13**, 291-318 (1979).
2. Kraemer, K. H. in *DNA Repair Processes*, 37-71 (Miami Symp. Spec., 1977).
3. McFarlin, D. E., Strober, W. & Waldmann, T. A. *Medicine* **61**, 281-314 (1972).
4. Sweet, H. O. *Mouse News Lett.* **65**, 27 (1981).
5. Green, M. C. *Genetic Variants and Strains of the Laboratory Mouse* (Fischer, New York, 1981).
6. Sidman, C. L., Shultz, L. D. & Unanue, E. R. *J. Immun.* **121**, 2392-2398 (1978).
7. Marshak-Rothstein, A. et al. *J. Immun.* **122**, 2491-2497 (1979).
8. Griselli, C. in *Immunodeficiency in Man and Animals*, 45-49 (Sinauer, Stamford, Connecticut, 1975).
9. Miller, T. E., Mackaness, G. B. & Lagrange, P. H. *J. natn. Cancer Inst.* **51**, 1669-1676.
10. Terplan, K. L. & Krauss, R. F. *Neurology* **19**, 446-454 (1969).
11. Reed, W. B., Epstein, W. L., Boder, E. & Sedwick, R. *J. Am. med. Ass.* **195**, 746-753 (1966).

Peripheral pathways are pioneered by an array of central and peripheral neurones in grasshopper embryos

Robert K. Ho & Corey S. Goodman

Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

During embryonic development, neuronal growth cones traverse long distances along stereotyped pathways to arrive at their appropriate targets. Since Harrison's initial observations¹, increasing evidence has supported the notion that axonal pathways are "pioneered" early in development when the distances are short and the terrain relatively simple; thus an axonal scaffold is erected upon which later neurones can navigate. Grasshopper embryos have emerged as an attractive preparation in which to study pathfinding because the nervous system is relatively simple, and the individual identified neurones in the periphery and central nervous system (CNS) of the embryo are large and highly accessible. How are the first axonal pathways established between the periphery and CNS? Bate² has described a single pair of cells in the limb buds, and two pairs in the antennae, which were thought to establish the very first peripheral pathways over distances of several hundred micrometres from the periphery to CNS; he called these pairs of cells 'pioneer neurones'. Keshishian³ confirmed this finding and also described⁴ a second pair of pioneers in the limb buds. We report here a further analysis of peripheral pathfinding in the grasshopper embryo using a monoclonal antibody, called I-5, which stains the cells responsible for establishing the first pathways^{5,6}. In contrast to the earlier reports, we find a large array of peripheral cells responsible for establishing the first pathways, with both pathfinders and landmark cells spaced at short intervals along the route. Furthermore, growth cones from central motoneurons make important contributions to the first peripheral pathways⁶.

Monoclonal antibodies against the grasshopper nervous system have been made with the goal of obtaining cell markers which label specific subsets of cells early in embryogenesis^{5,6}. In early embryos, one of these antibodies, called I-5, stains a specific pattern of neurones and mesodermal cells^{5,6}, including the pairs of peripheral pioneer neurones originally described by Bate² and Keshishian^{3,4}. Interestingly, the I-5 antibody also reveals an array of previously unseen cells and growth cones that are intimately involved in pioneering the first peripheral axonal pathways. In this paper we have studied these cells using the I-5 monoclonal antibody^{5,6}, observations with Nomarski

optics^{7,8}, intracellular injections of the fluorescent dye Lucifer Yellow⁹, and an anti-Lucifer Yellow antibody (anti-LY)¹⁰.

The I-5 antibody stains the pairs of pioneer neurones in both the antennae² and limb buds²⁻⁴ shortly after their final cell division, and before axonogenesis (Fig. 1a); the antibody also stains their growth cones and axons (Fig. 1b,c). In the antenna (Fig. 2), two pairs of neurones [first the ventral pioneers (VP) and then the dorsal pioneers (DP)] appear at the distal tip and extend growth cones towards the CNS, thus pioneering the two axonal pathways within the antenna. At the same time, a single base pioneer (BP) appears at the base of the antenna and its growth cone is the first peripheral process to reach the CNS. Simultaneously, a growth cone from a motoneurone originating in the CNS extends outwards towards the base of the antenna, running parallel, although not fasciculating upon, the BP's axon; the motoneurone growth cone stops at the lateral edge of the antennal base and in time innervates a muscle that develops at this location. Thus, the pathway between the antennal base and the CNS is complete by the time the growth cones from the more distal VP and DP neurones arrive and follow the BP pathway into the CNS. Interestingly, when these cells first appear inside the ectodermal epithelium, they are spaced at

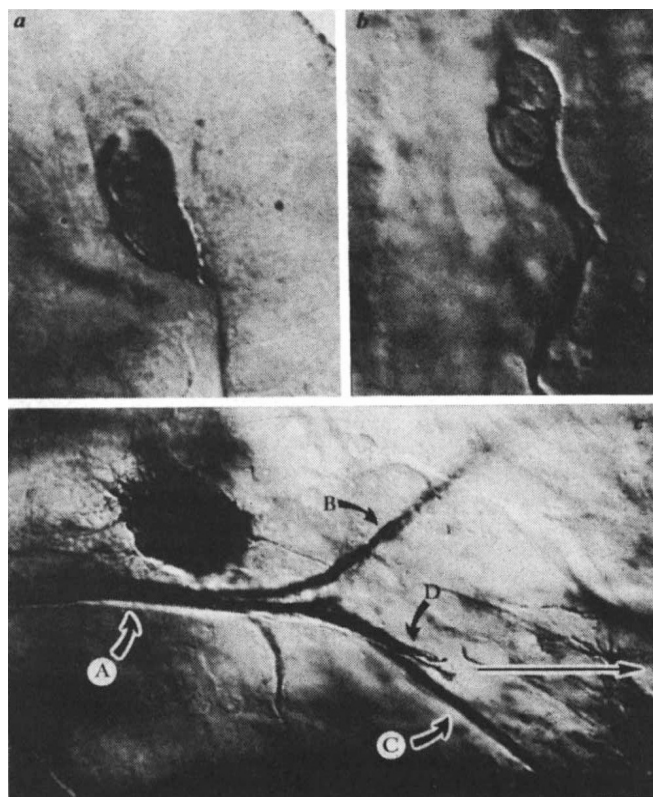


Fig. 1 Photomicrographs of limb buds (A, C) and antennae (B) showing staining of cells with I-5 monoclonal antibody; the growth cones of these cells are involved in pioneering the peripheral axonal pathways in the grasshopper embryo. The embryo is dissected out of the eggshell, lightly fixed in 2% paraformaldehyde, and incubated overnight in I-5 supernatant with 0.2% saponin. The antibody is visualized using an HRP-conjugated second antibody. *a*, Staining of the pair of 1E cells in the limb bud just before they initiate growth cones. The 1E cells are just within the ectodermal epithelium at the tip of the limb bud. *b*, Staining of the ventral pioneers (VP) in the antennae after they have initiated axons. Their growth cones (not shown) have just arrived at the cell body of the base pioneer (BP). *c*, Staining of the limb bud showing the intersection of the A, B, and C pathways (established by the 1A, 1B, and 1C cells, respectively), and the growth cones of the early motoneurons (the β motoneurons) as they grow distally and establish the D pathway (long arrow) (see Figs 3, 4). Large cell above A pathway is a mesodermal cell also stained by the I-5 antibody. Cell body diameters in *a* and *b* are $\sim 15 \mu\text{m}$; large mesodermal cell in *c* is $\sim 30 \mu\text{m}$ in diameter.

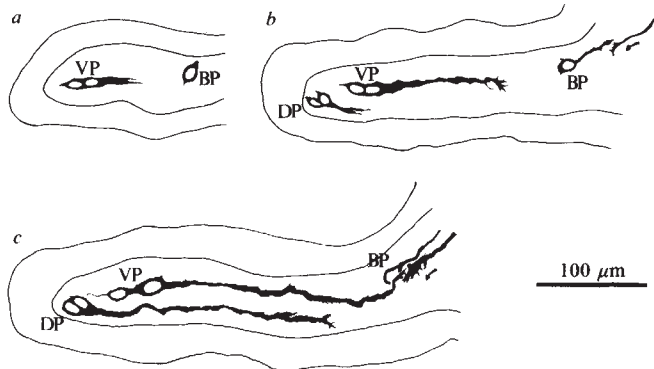


Fig. 2 Drawings of the antennae of the grasshopper embryo at successively older stages of development, showing the pioneering of the peripheral axonal pathways by the growth cones from peripheral and central (arrows) cells. Bar, 100 μ m.

relatively short distances from one another; the VP neurones are less than 100 μ m from the BP neurone, and the BP neurone is less than 100 μ m from the CNS. Lucifer Yellow injections (and visualization with anti-Lucifer Yellow) reveal that the cells before axonogenesis and the growth cones afterwards radiate long, thin filopodia that typically reach over 50 μ m and sometimes for up to 100 μ m (ref. 10). Thus, the cells may be within filopodial grasp at, or shortly after, axonogenesis.

A similar, albeit more complex, series of events takes place in developing limb buds (Figs 3, 4). The first growth cones in the limb bud arise from the 1B cells (Bate's original 'pioneers'). The 1B growth cones extend proximally towards the CNS along the basement membrane on the inside of the ectodermal epithelium of the limb bud, growing past where the 2B cell appears and towards the 3B cell (the B pathway, Fig. 4). However, the 1B growth cones make a characteristic turn about the time they reach the 3B cell (although they need not reach this cell); they always, however, turn towards the 1A cells which are about 50 μ m away and within filopodial grasp (Fig. 3). Upon reaching the 1A cell bodies, the 1B growth cones then continue extending towards the CNS. One source of guidance for the 1B growth cones is likely to be the 1A cells (as 'landmark' cells). In the pro- and mesothoracic limb buds, the 1A cells send growth cones towards the CNS before the 1B growth cones arrive at their cell bodies; thus the 1A cells pioneer the first pathway (the A pathway, Fig. 4) between the base of the limb bud and the CNS (Fig. 3a). The timing is different, however, in the metathoracic limb bud. Here the 1A cells do not pioneer the first pathway to the CNS; rather, the 1B cells grow towards them and then continue growing towards the CNS, followed shortly thereafter by the 1A growth cones.

A second pathway is simultaneously made by growth cones of motoneurons originating in the CNS; these growth cones do not fasciculate with the 1A or 1B axons but rather form a separate yet parallel pathway between the base and the CNS. As development proceeds, the A pathway becomes nerve 5 whereas the second parallel pathway becomes nerve 3.

The C pathway is made by the pair of 1C neurones. Much of the D pathway is pioneered by growth cones of motoneurons originating in the CNS which extend out along the A pathway shortly after it is established; instead of growing along the B or C pathways, these motoneurons make a new pathway as

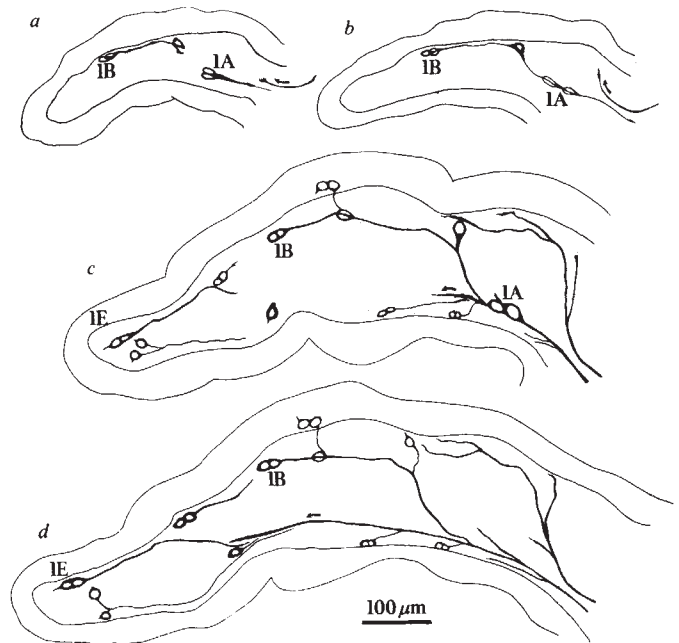


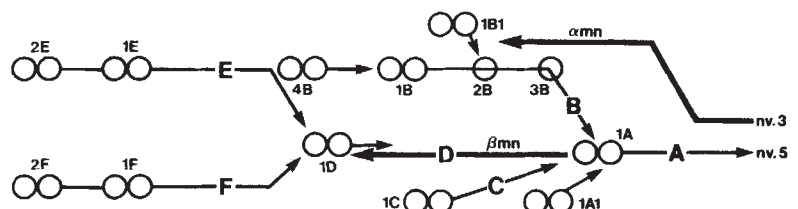
Fig. 3 Drawings of the mesothoracic limb bud of the grasshopper embryo at successively older stages of development, showing the pioneering of the peripheral axonal pathways by the growth cones from peripheral and central (arrows) cells. For clarity, only a few cells are named; the complete identities are shown in Fig. 4.

they grow distally (Figs 1, 3, 4). These motoneuron growth cones meet those from the 1D cells and thus the D pathway is completed. The E and F pathways are pioneered by pairs of peripheral cells in a similar fashion simultaneous to the pioneering of the D pathway (Figs 3, 4) [the 1E cells are Keshishian's⁴ second pair of pioneers]. Just as in the antenna, the distances between the peripheral neurones in the limb bud are relatively short (often 50 to 100 μ m) when they first extend growth cones. Here too the cells have extensive filopodia and may be in direct contact at, or shortly after, axonogenesis.

Four major conclusions are suggested by our results. First, peripheral pathways are pioneered by growth cones from central as well as peripheral neurones. For example, nerve 3 and the D pathway of nerve 5 are pioneered by motoneurons.

Second, there are not just two pairs of cells involved in pioneering the peripheral pathways in limb buds, but rather a large array of cells. In sheer number, this array of cells (Fig. 4) at first glance appears to hopelessly increase the complexity of the problem. Yet their unveiling dramatically simplifies the picture and points us towards specific experiments. The finding that the cells are so close and no one cell must alone traverse a large expanse of axonless territory, shifts our attention from long distance guidance mechanisms (e.g. diffusible gradients) to local interactions between identified cells over short distances. In many cases, cells may be in direct filopodial contact with 'landmark' cells (for example the 1A cells appear to be stepping stones [2] for the 1B growth cones); in other cases, specific mechanical substrates (e.g. the tendon followed by the 1E cells, see below) may take the growth cones to their next neuronal contact. The finding of landmark cells in the periphery is similar to our finding of landmark cells in the CNS¹².

Fig. 4 Schematic diagram showing the identities of cells responsible for pioneering the pathways A to F of nerve 5, and the first pathway of nerve 3 in the pro- and mesothoracic limb buds. In the metathoracic limb bud, the 1A cells do not pioneer the A pathway, but rather the 1B growth cones use the 1A cells as landmarks and then continue growing towards the CNS.



Third, growth cones in the periphery are not simply guided by passive mechanisms. Active guidance mechanisms are suggested because certain growth cones are confronted with several possible choices and make cell-specific decisions of which way to go. For example, the 1E growth cones extend along the epithelium until they near the 4B cell bodies. Rather than climb onto the B pathway, they make a turn and grow along a tendon to the 1D cells, thus completing the E pathway. In a similar way, early motoneurone growth cones are confronted with the B and C pathways, yet choose to grow laterally and pioneer a new D pathway. The finding of specific choices made by growth cones in the periphery is similar to the divergent and specific choices made by identified growth cones in the CNS¹¹⁻¹⁴.

Finally, we must re-evaluate the concept of 'pioneer neurones'. Bate and Goodman have shown that other peripheral nerves (such as the intersegmental nerve and the median nerve) are pioneered by motoneurone growth cones from neuroblast progeny in the CNS (manuscript in preparation). Furthermore, pathways in the CNS are pioneered by

progeny from two different classes of neuronal precursors—the midline precursors and neuroblasts^{12,15}. Thus, there is no specialized class of precursor cell which gives rise to cells having the unique role of establishing axonal pathways throughout the grasshopper embryo. Rather, neurones of diverse ectodermal origin appear early in embryogenesis and pioneer pathways when the distances are short and the terrain relatively simple. This finding shifts the emphasis away from the special role played by 'pioneers', and focuses our attention on the questions of what cell-to-cell interactions and substrate features guide the first peripheral growth cones. These questions can be answered by manipulating the pioneer neurones, landmark cells, and ectodermal epithelium in grasshopper embryo cultures. In this way, we hope to establish by what mechanisms these 'pioneering' growth cones are guided across short distances of axon-less territory.

We thank Susannah Chang for the I-5 antibody and Paul Taghert and Michael Bastiani for the anti-LY antibody. This work was supported by grants from the McKnight Foundation, NSF, and Sloan Foundation to C.S.G.

Received 30 November 1981; accepted 2 April 1982.

1. Harrison, R. G. *J. exp. Zool.* **9**, 787-846 (1910).
2. Bate, C. M. *Nature* **260**, 54-56 (1976).
3. Keshishian, H. *Dev. Biol.* **80**, 388-397 (1980).
4. Keshishian, H. & Bentley, D. *Soc. Neurosci.* **7**, 347 (1981).
5. Chang, S., Ho, R., Raper, J. A. & Goodman, C. S. *Soc. Neurosci.* **7**, 347 (1981).
6. Ho, R., Chang, S. & Goodman, C. S. *Soc. Neurosci.* **7**, 348 (1981).
7. Goodman, C. S. & Spitzer, N. C. *Nature* **280**, 208-214 (1979).

8. Goodman, C. S., O'Shea, M., McCaman, R. E. & Spitzer, N. C. *Science* **204**, 1219-1222 (1979).
9. Stewart, W. W. *Cell* **14**, 741-759 (1978).
10. Taghert, P., Bastiani, M., Ho, R. K. & Goodman, C. S. *Cell* (submitted).
11. Raper, J. A. & Goodman, C. S. *Soc. Neurosci.* **7**, 347 (1981).
12. Goodman, C. S., Raper, J. A., Ho, R. & Chang, S. *40th Symp. Soc. dev. Biol.* (in the press).
13. Raper, J. A., Bastiani, M. & Goodman, C. S. *J. Neurosci.* (in the press).
14. Raper, J. A., Bastiani, M. & Goodman, C. S. *J. Neurosci.* (in the press).
15. Bate, C. M. & Grunewald, E. B. *J. Embryol. Exp. Morphol.* **61**, 317-330 (1981).

Electrophysiology of mammalian thalamic neurones *in vitro*

Rodolfo Llinás & Henrik Jahnsen

Department of Physiology and Biophysics,
New York University Medical Center,
550 First Avenue, New York, New York 10016, USA

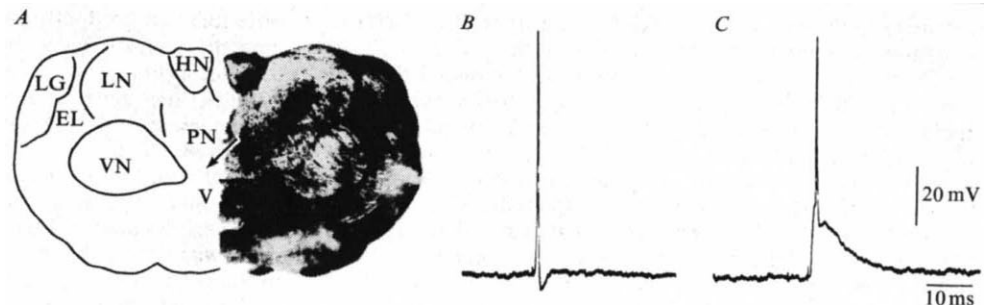
Although much is known about the morphology and physiology of thalamic neurones¹, no information is available regarding the ionic basis for the electrical excitability of these cells. Furthermore, possible differences in the electrical properties of the principal nerve cells in the various thalamic groups have not been studied in sufficient detail to determine whether the thalamus is, electrophysiologically, a uniform set of neuronal elements. Here we present evidence that thalamic neurones have voltage-sensitive ionic conductances capable of generating two distinct functional states—a repetitive and a burst-firing mode. The neurones are switched from one state to the other by membrane potential changes, each state being dominated by different voltage-dependent ionic conductances. At membrane potentials more positive than -60 mV, the neurones respond to a depolarization with repetitive firing via Na⁺-dependent action potentials, whereas at potentials more negative than -65 mV, depolarization of the cell results in short bursts of action potentials via an inactivating all-or-none Ca²⁺-dependent spike. This property, present in all the neurones comprising the different thalamic nuclei, serves as the basis for their oscillatory properties. Particularly, the inactivating Ca²⁺ conductance represents the ionic basis for the post-anodal excitation, the mechanism most probably responsible for the alpha rhythm.

The study was performed on 400-μm-thick thalamic slices obtained by coronal section of guinea pig brain using a vibratome cutting device (Oxford G501 Vibratome). Typically, six sections included most of the thalamus. The sections were placed in a special lucite chamber under a flowing mammalian Ringer's solution². Electrophysiological analysis was performed, under direct microscopic observation, on neurones of the anterior, medial and lateral thalamic nuclear groups, including the lateral geniculate nucleus and neurones in the lamina

interna. The techniques for sectioning and maintenance of the slice have been described elsewhere². Intracellular recordings were made using potassium citrate-filled micropipettes having a d.c. resistance of 50-90 MΩ. Cells were activated either antidromically or synaptically by means of bipolar stimulating electrodes located on the midline region of the thalamic mass (see arrow in Fig. 1A) or directly through the microelectrode using a bridge amplifier. Neurones were identified by their antidromic and synaptic activation, and their thalamic site was determined by direct observation. Although single somata were often not visible, the exact outline of the groups of nuclei could be easily determined (see Fig. 1A). Intracellular recordings from these neurones indicated that they could survive well for >8 h after the slicing procedure. The stimulus amplitude required for the antidromic activation of these cells (Fig. 1B) was often lower than that for the activation of synaptic inputs (Figs 1C, 2E, F). Considering the thickness of the slices, we suspect that only monosynaptic connectivity can be reliably studied in this material.

The electrical properties of the cells were examined by direct stimulation via the recording electrode. As shown in Fig. 2A, thalamic cells may be characterized by two distinct types of electrical response. When the resting potential was more positive than -60 mV, thalamic neurones fired repetitively at increasing frequencies with increasing amplitude of direct depolarization. A plot of the instantaneous adapted firing frequency, *f* (1 s pulse duration) against current injection (*I*) is shown in Fig. 2C, which illustrates the graded nature of the repetitive firing properties of these cells. However, in contrast to other central neurones³, no primary firing range was observed: rather, the neurones fired with a typical minimal frequency of ~10 impulses s⁻¹ and reached firing frequencies of 165 impulses s⁻¹. At this membrane potential level, we observed no after-depolarization comparable to that seen in motoneurones⁴, and Purkinje², inferior olivary⁵ and hippocampal⁶ cells *in vitro*. On the other hand, when the cells had resting potentials more negative than -60 mV, or were hyperpolarized by direct inward current injections, rather than the continuous firing, a single burst of action potentials was observed after a step depolarization (see Fig. 2A, B). The lower record in Fig. 2A demonstrates a passive membrane depolarization produced by a square current pulse that is subthreshold for spike initiation. When, as shown in the upper record in Fig. 2A, the same

Fig. 1 Neuronal localization in thalamic nuclei. **A**, photomicrograph of diencephalic slice (right) and diagram illustrating the location of thalamic nuclei (left). EL, external medullary lamina; HN, habenular nucleus; LG, lateral geniculate nuclei; LN, lateral nucleus; PN, parafascicular nucleus; VN, ventral thalamic nucleus; V, third ventricle. Arrow indicates the location of the bipolar stimulating electrode. **B**, antidromic invasion of VN cell after midline stimulation. **C**, synaptic activation from the same locus after increase in stimulus strength. The voltage and time calibration are indicated.



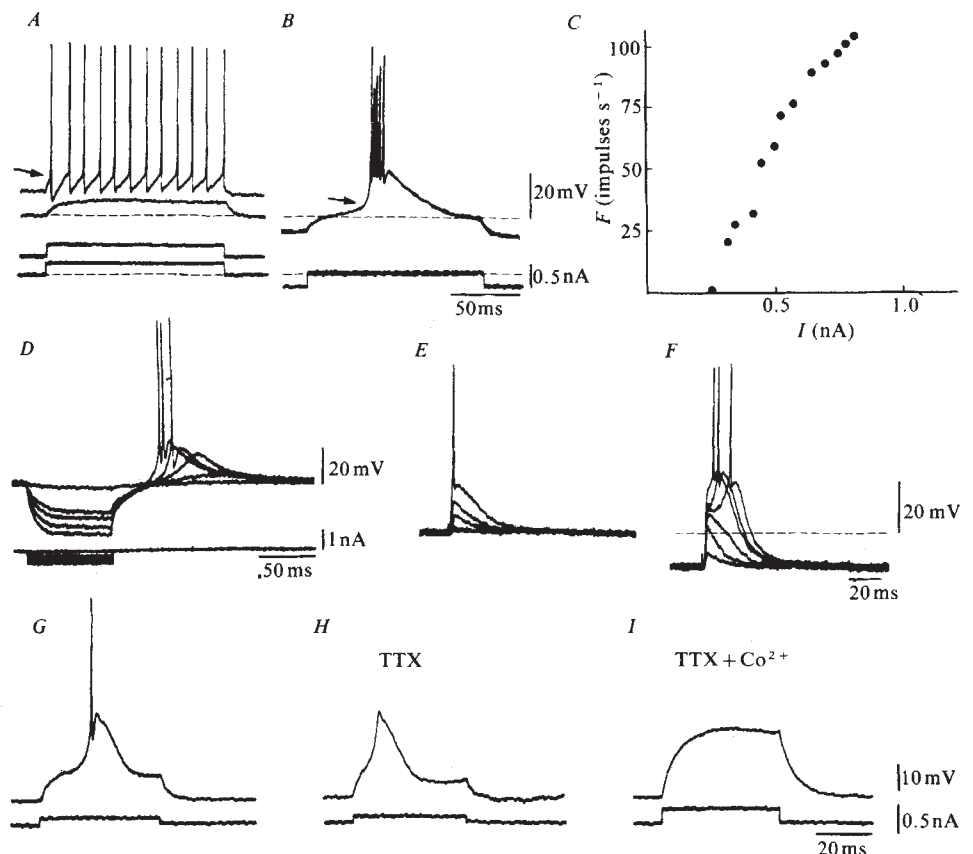
current pulse was superimposed on a d.c. depolarization, the cell fired repetitively. This is expected if the sum of the two depolarizations reaches the firing level of the cell. However, when the membrane potential was hyperpolarized from rest, another type of electroresponsiveness was observed. Indeed, the excitability of the cell was increased by membrane hyperpolarization. Thus, single burst responses having an abrupt onset and a smooth falling phase were obtained by current pulses as in Fig. 2A when the cell was previously hyperpolarized by 7 mV (Fig. 2B). Note the marked difference between the firing level for the tonic firing (arrow in Fig. 2A) and that of the burst response (arrow in Fig. 2B). The burst response had two distinct components: (1) a slowly rising all-or-none depolarization, the amplitude of which is directly related to the level of membrane hyperpolarization; and (2) one or more fast spikes activated from the slow depolarization.

Such a burst was also observed as a rebound response at the break of a hyperpolarizing current pulse (Fig. 2D). In these conditions the two components could be easily separated. In fact, with low-amplitude hyperpolarizing pulses only the slow component was observed, emphasizing the graded nature of

the amplitude of this response. With larger hyperpolarizations the slow component reached the threshold for fast spike activation. Note that the hyperpolarization had a slow recovery time suggesting the presence of an 'early Na conductance'. The increased excitability that occurred with hyperpolarization was also demonstrated after synaptic activation. Figure 2E shows that smoothly graded excitatory synaptic potentials which reached the firing level were obtained in a thalamic neurone having a resting potential of -60 mV after midline thalamic stimulation of increasing amplitude. When the neurone was hyperpolarized by 15 mV (Fig. 2F), the same synaptic activation generated all-or-none slow depolarization, which in turn activated fast spikes. Note once again the different firing levels for the activation of the slow response and the fast action potentials.

These two modes of electroresponsiveness were found to have different ionic mechanisms as demonstrated by the addition of specific ionic channel blockers to the bathing solution. The fast action potentials, which seem to arise from the soma and axon areas (judging from the extracellular negativity they generate) can be blocked by tetrodotoxin (TTX) indicating that they are sodium-dependent⁸. As shown in Fig. 2G, a step

Fig. 2 Electrophysiological and pharmacological properties of thalamic neurones. **A**, a subthreshold depolarization of the cell at resting level (broken line) produces, after a d.c. depolarization, repetitive firing of the cell during the same current pulse. **B**, after d.c. hyperpolarization, similar current pulses as in **A** produce a single high frequency spike burst. **C**, plot of the instantaneous adapted frequency (f) of the repetitive firing of the cells as seen in **A** for different levels of current injection (I). **D**, rebound burst response after hyperpolarizing pulses of different amplitudes. **E**, in another cell, excitatory synaptic potentials obtained from rest show their graded character and the firing level of the neurone. **F**, synaptic potentials of similar amplitude as in **E** produce, following hyperpolarization, long-lasting all-or-none responses on which fast spikes are generated at the same firing level as in **E**. **G**, slow all-or-none response generating fast spike, from a cell slightly hyperpolarized from rest potential. **H**, after blockage of Na^+ conductance with TTX, the fast action potential is blocked but the slow response remains unchanged. **I**, addition of CoCl_2 to the bathing solution abolishes completely the slow response seen in **H**, even when the current pulse is increased in amplitude by 2.5 times.



depolarization from a holding level of -65 mV produced a large all-or-none depolarization and a secondary fast spike. Addition of TTX to the bath at $1 \mu\text{g ml}^{-1}$ completely blocked the fast action potential, leaving the underlying all-or-none slow depolarization unaltered (Fig. 2H). By contrast, addition of Ca^{2+} -conductance blockers such as CoCl_2 (ref. 9) or CdCl_2 (ref. 10) to the bathing solution at a concentration of 1 mM, or substitution of MnCl_2 (3.5 mM) for CaCl_2 , abolished this all-or-none response (Fig. 2I). These results indicate that the slow spike is generated by a voltage-dependent change in Ca^{2+} conductance¹¹. Furthermore, the nature of this response is similar to the low-threshold Ca^{2+} -dependent spikes described in the inferior olive^{5,12}.

A more detailed study of the switching between a Ca^{2+} -dominated and a Na^{+} -dominated electroresponsiveness revealed that this change occurs in almost an all-or-none manner. Figure 3 shows that when the excitability of a cell was challenged with a series of brief depolarizing pulses (of sufficient amplitude to fire the cell), and simultaneously the membrane

potential was gradually decreased by a slow ramp depolarizing current, the response to the pulse changed from the burst mode to tonic firing. In this case the bursts had a duration of 30–130 ms, and generated a set of fast action potentials at a frequency of 150–320 impulses s^{-1} , depending on the level of membrane hyperpolarization. As the membrane was slowly depolarized, the test pulses produced, rather than the burst, a continuous repetitive spiking for the duration of the pulse (compare *b* and *c* in Fig. 3B). The frequency of this repetitive firing was related to the total level of depolarization as indicated in the *f-I* plot in Fig. 2C.

The mechanism for the switching between these two excitability states was clarified by further studying the properties of the Ca^{2+} -dependent electroresponsiveness. As shown in Fig. 2D, the amplitude and the rate of rise of this Ca^{2+} -dependent response varies with membrane potential in a very steep manner. In addition, as in the case of the inferior olivary cells, these responses completely inactivate at membrane potential levels more positive than -60 mV. This voltage-dependent inactivation explains both the presence of burst responses when the cell is hyperpolarized and the rebound activation of these cells. Furthermore, as observed in inferior olivary neurones⁵, these rebound calcium spikes show a long refraction, indicating that the kinetics of recovery from the inactivated state are slow. This slow recovery also explains why such responses are unitary as their activation is followed by a rather prolonged refractory period of 80–150 ms.

In conclusion, these findings indicate that thalamic neurones have special voltage-sensitive ionic conductance properties which allow them to change from a phasic bursting response (followed by neuronal silence) to graded repetitive firing, and this switching is modulated by membrane potential. When the cells have a resting level more negative than -60 mV, a low threshold calcium response is obtained. The other integrative state is characterized by tonic firing brought about by depolarization of the cell to levels more positive than -60 mV. At this membrane potential level, the calcium conductance which generates the slow response is inactivated and the cell fires repetitively. Although the precise distribution of these conductances over the soma-dendritic area of these neurones has not been determined, by analogy with inferior olivary cells^{5,12}, we suggest that the Ca^{2+} -deactivated response probably resides at the somatic level, as does the Na^{+} -dependent action potential.

These electrical properties further explain many of the electroresponsive properties observed in thalamic recordings *in vivo* where such switching of tonic to phasic firing has in fact been described^{13,14}. Moreover, in conjunction with the recurrent inhibition observed *in vivo*, this newly described Ca^{2+} conductance represents the functional basis of the so-called 'post-anodal exaltation' underlying the recruiting responses¹² at cortical level and most probably the α rhythm¹⁵.

This research was supported by USPHS grant NS13742 from the National Institute of Neurological and Communicative Disorders and Stroke. H.J. was supported in part by grants from the University of Copenhagen, the Danish MRC and the Weimann Foundation, and by an Albert Cass Travelling Fellowship.

Received 21 January; accepted 29 March 1982.

1. Shepherd, G. D. *The Synaptic Organization of the Brain*, 2nd edn (Oxford University Press, New York & Oxford, 1979).
2. Llinas, R. & Sugimori, M. *J. Physiol., Lond.* **305**, 171–195 (1980).
3. Granit, R. *Mechanisms Regulating the Discharge of Motoneurons* (Charles C. Thomas, Illinois, 1972).
4. Fulton, B. P., Miledi, R. & Takahashi, T. *Proc. R. Soc. B* **208**, 115–120 (1980).
5. Llinas, R. & Yarom, Y. *J. Physiol., Lond.* **315**, 549–567 (1981).
6. Schwartzkroin, P. A. *Brain Res.* **85**, 423–436 (1975).
7. Connor, J. A. & Stevens, C. F. *J. Physiol., Lond.* **213**, 21–30 (1971).
8. Narahashi, T., Moore, J. W. & Scott, W. R. *J. gen. Physiol.* **47**, 965–974 (1964).
9. Baker, P. F., Hodgkin, A. L. & Ridgway, E. G. *J. Physiol., Lond.* **218**, 709–755 (1971).
10. Kostyuk, P. G. & Krishtal, O. A. *J. Physiol., Lond.* **270**, 545–568 (1977).
11. Hagiwara, S. *Adv. Biophys.* **4**, 71–102 (1973).
12. Llinas, R. & Yarom, Y. *J. Physiol., Lond.* **315**, 569–581 (1981).
13. Purpura, D. P. & Cohen, B. *J. Neurophysiol.* **25**, 621–635 (1962).
14. Maendly, R. et al. *J. Neurophysiol.* **46**, 901–917 (1981).
15. Andersen, P., Eccles, J. C. & Sears, T. A. *J. Physiol., Lond.* **174**, 370–399 (1964).

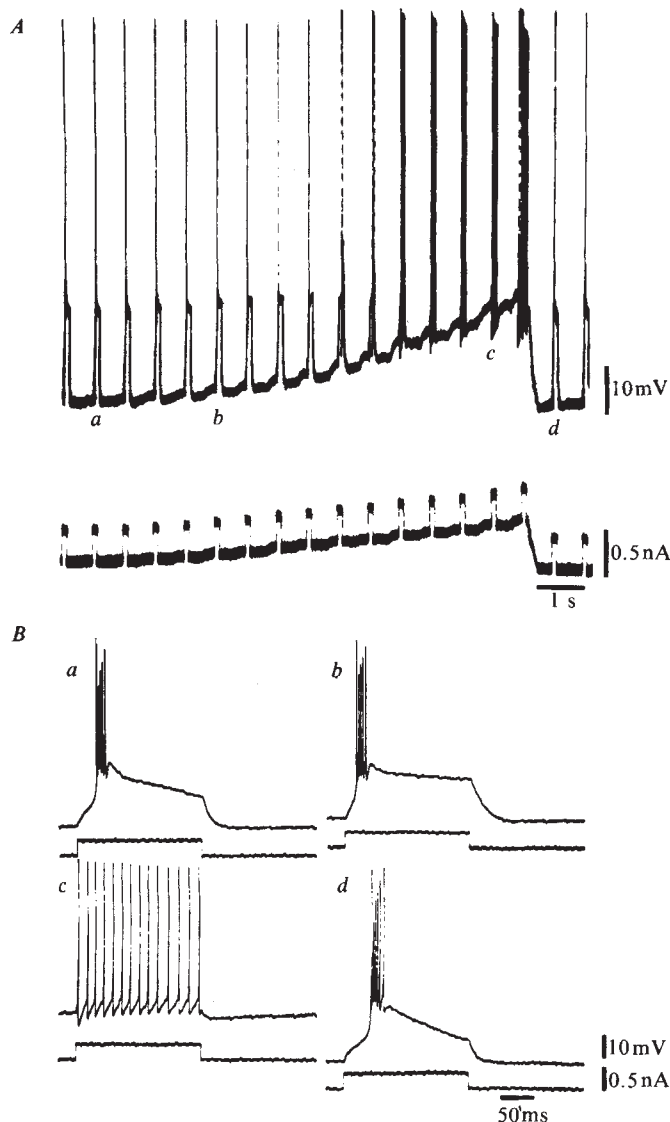


Fig. 3 Voltage-dependent burst-to-tonic switching of thalamic cell activity. *A*, response of a thalamic cell after short current pulses delivered from a slowly rising ramp depolarization pulse. The cell switched abruptly from a burst response to tonic firing as the d.c. potential decreased by ~ 10 mV from the initial value. *B*, records obtained at a higher sweep speed at the times indicated by *a* to *d* in *A*. Note the transition from burst response (*a* and *b*) to tonic response (*c*), followed by the abrupt return to a burst response (*d*).

Evidence for the presence of S-100 protein in the glial component of the human enteric nervous system

Gian-Luca Ferri*, Lesley Probert*,
Domenico Cocchia†, Fabrizio Michetti†,
Paul J. Marangos‡ & Julia M. Polak*

* Department of Histochemistry, Royal Postgraduate Medical School, Hammersmith Hospital, London W12 0HS, UK

† Department of Anatomy, Università Cattolica, 00168 Roma, Italy

‡ Clinical Psychobiology Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20205, USA

Recently, the complex organization of the enteric nervous system and its independent role in the integrative control of digestive functions have been fully recognized. However, little is known about the non-neuronal elements present in close morphological and, presumably, functional relationship with the enteric neurones. Although 'enteric glial cells' in gut ganglia and Schwann cells accompanying nerve fibres throughout the gastrointestinal wall can easily be recognized by electron microscopy on the basis of their ultrastructural features¹⁻³, the lack of a suitable common marker for the visualization of these non-neuronal components of the enteric nervous system has hampered investigation. S-100 is a Ca^{2+} -binding protein⁴ originally isolated from the brain⁵, where it is localized mainly in glia^{6,7}. We demonstrate here, using immunohistochemistry and electron immunocytochemistry, that both enteric glial cells and Schwann cells of the human gut contain densely immunoreactive S-100. This protein can therefore be regarded as a common marker for the glial components of the enteric nervous system.

The distribution of the S-100 protein in the enteric nervous system was studied by immunocytochemistry in tissue samples obtained from uninvolved areas of segments of human gut resected for carcinoma (stomach, ileum and colon; $n = 15$). Throughout the entire length of the gastrointestinal tract, S-100-immunoreactivity was detected in the ganglia of both the submucous and the myenteric plexus, in slender cells having numerous projections partially surrounding enteric neurones (Fig. 1a), presumably glial cells. In addition, numerous elongated, dendritic cellular elements were immunostained in all non-epithelial layers of the wall (Fig. 2); these were considered to be Schwann cells.

To confirm that in enteric ganglia S-100 is selectively localized in glial cells, we compared the distribution of this antigen with that of neurone-specific enolase (NSE), an isomer of the glycolytic enzyme enolase that is specifically localized in neurones of the central nervous system (CNS)^{7,8} and has also been demonstrated in gut neurones⁹. Figure 1 shows that S-100 was localized only in glial cells (a), while neurones, identified by NSE immunostaining (b), were completely negative.

Ultrastructurally identifiable Schwann cells accompany the bundles of nerve fibres running through all layers of the gut wall, forming an incomplete sheath around them². To confirm that the S-100-immunoreactive elements found outside the enteric ganglia were Schwann cells, we studied their three-dimensional structure and distribution in stretch preparations of either submucosa or lamina propria separated by microdissection, then immunostained and examined whole¹⁰. In both of these layers, a network of S-100-immunoreactive elements was demonstrated in the internodal strands running between enteric ganglia (Fig. 3a), around blood vessels, and in a dense mesh of interconnecting cells both at the level of the muscularis

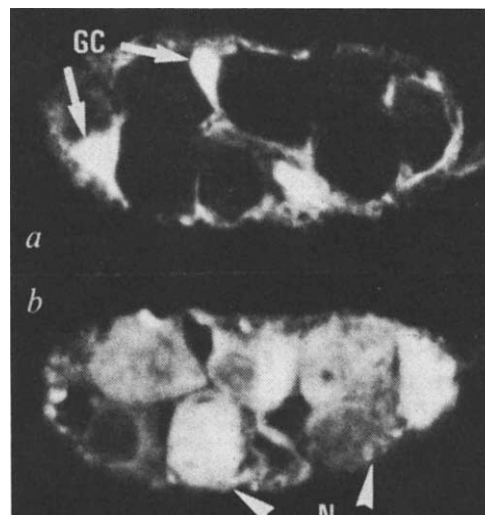


Fig. 1 Immunofluorescent detection of S-100 (a) and NSE (b) in two serial (4 μm) sections of a submucous ganglion in the human colon. Tissues were fixed in *p*-benzoquinone solution²¹ (0.4% in phosphate-buffered saline, PBS) for 2 h, washed overnight in PBS containing 7% sucrose, frozen in melting Arcton and cut with a cryostat. Sections were stained by indirect immunofluorescence using anti-S-100 antiserum diluted 1:500, and anti-NSE antiserum diluted 1:800 in PBS (overnight incubation), followed by fluorescein isothiocyanate-conjugated goat anti-rabbit IgG diluted 1:200 (1 h). Specific rabbit antisera raised against ox S-100¹⁴ and rat NSE⁸, fully cross-reacting with the human S-100 and NSE, were used. Controls included the use of antibodies preabsorbed with 20 $\mu\text{g ml}^{-1}$ of ox S-100 and 5 nmol ml^{-1} of rat or human NSE. S-100 immunoreactivity (a) was confined to the slender and elongated glial cells (GC) partially surrounding neurones (N), which were completely negative for this antigen and were revealed (b) by NSE immunostaining. $\times 700$.

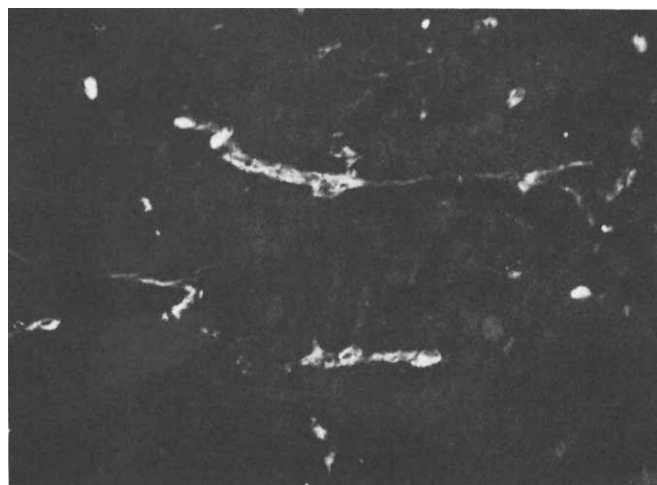


Fig. 2 S-100 immunofluorescence (fixation and immunostaining as described in Fig. 1 legend) in the muscle layer of the human colon. Numerous slender cells having several long, branching processes can be observed. Similar elements, with the morphological appearance of Schwann cells, were demonstrated in all non-epithelial layers of the gut wall (10 μm section, $\times 120$).

mucosae (Fig. 3b) and in the lamina propria up to close contact with the basement membrane of the epithelium. The distribution of S-100-immunoreactivity seems therefore to parallel the general distribution of the enteric innervation.

Glial cells in enteric ganglia have distinctive ultrastructural features resembling those of astrocytes in the CNS³. Typical

glial cells identified by electron microscopy in enteric ganglia contain S-100-immunoreactivity, while enteric neurones are negative (Fig. 4a). In addition, in all gut layers, positive immunostaining for S-100 was demonstrated in Schwann cells partially surrounding bundles of nerve fibres (Fig. 4b). In the present experiments, only the cytoplasmic localization was consistently demonstrated in enteric glia, even though a nuclear localization of this protein has been documented in the CNS¹¹⁻¹³.

Although S-100 has recently been demonstrated in various locations outside the nervous system, including stellate cells in the adenohypophysis^{14,15}, satellite cells in the adrenal medulla¹⁶, melanocytes and Langerhans cells in the skin¹⁷ and chondrocytes¹⁸, there is general agreement on its prevalent localization in the CNS in glia^{6,7}. The presence of this brain protein in the enteric nervous system agrees with the numerous homologies reported between the gut 'mini brain' and the CNS².

Another brain protein, the glial fibrillary acidic protein, has recently been demonstrated in enteric glial cells¹⁹. However, the absence of this protein from Schwann cells¹⁹ precludes its use as a common marker of the non-neuronal components of the enteric nervous system.

In conclusion, our results indicate that the S-100 protein can be reliably used to demonstrate the glial elements of the enteric nervous system in man. This will have immediate relevance to the study of disease conditions affecting the gastrointestinal

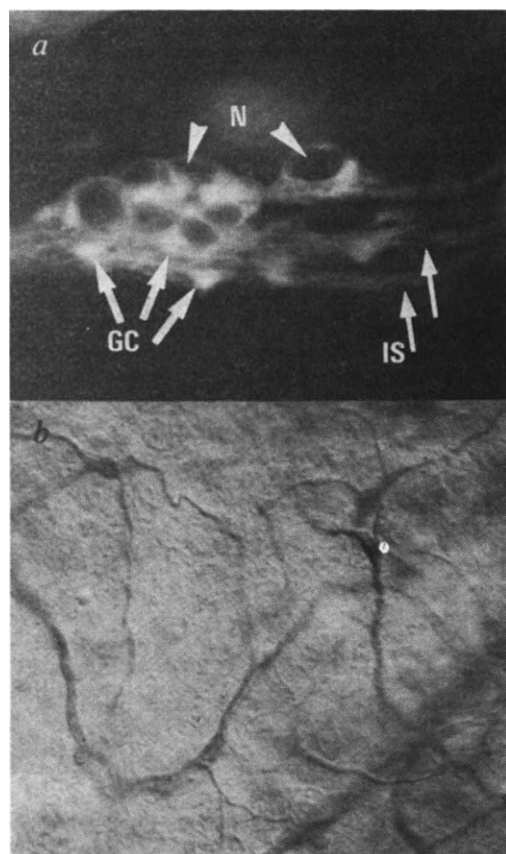


Fig. 3 S-100 immunoreactivity in whole-mount preparations of submucosa¹⁰. This layer was separated from samples of ileum and colon by microdissection, fixed in *p*-benzoquinone (see Fig. 1 legend), dehydrated through graded alcohols, cleared in xylene, rehydrated and immunostained *in toto* by immunofluorescence or by the peroxidase-antiperoxidase technique²². Prolonged incubations (16–36 h) and washings (6–8 h) were used. S-100 immunostaining revealed the general distribution of glial elements in the whole enteric innervation; *a*, in glial cells (GC) partially circumscribing unstained neurones (N) in the ganglia, and in Schwann cells in internodal strands (IS); and *b*, in a rich network of Schwann cells at the level of the muscularis mucosae. (*a*, $\times 375$; *b*, $\times 280$.)

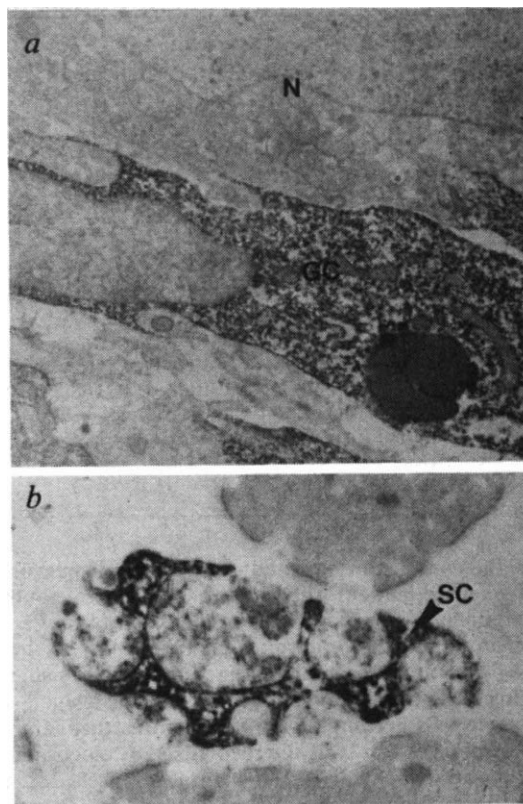


Fig. 4 Electron immunocytochemistry for S-100. Tissues were fixed in 4% paraformaldehyde in 0.1 M phosphate buffer for 3 h, washed overnight in PBS and cut at 100 μ m with a vibratome. Sections were stained by the peroxidase-antiperoxidase technique^{14,23} (same antibody dilution and controls as described for Fig. 1) using diaminobenzidine acid as coupler, then embedded in Araldite and cut at 60 nm for electron microscopy. In enteric ganglia (*a*), immunoreactivity is demonstrated in the cytoplasm of glial cells (GC), which partially circumscribe neurones that are completely unstained (N). In a nerve bundle in the muscle layer (*b*), projections of Schwann cells (SC) are in close contact with nerve fibres, which are unstained (ileum; *a*, $\times 10,000$; *b*, $\times 20,000$).

innervation, especially those in which an abnormal glial proliferation is suggested²⁰.

G.-L.F. is a research fellow from the Department of Medicine 1, University of Bologna, Italy, receiving financial support from the Accademia Nazionale dei Lincei, L.P. is a MRC student.

Received 5 February; accepted 7 April 1982.

1. Cook, R. D. & Burnstock, G. *J. Neurocytol.* **5**, 195–206 (1976).
2. Gabella, G. *Int. Rev. Cytol.* **59**, 129–193 (1979).
3. Gabella, G. *Neuroscience* **6**, 425–436 (1981).
4. Klee, C. B., Crouch, T. H. & Richman, P. G. *A Rev. Biochem.* **49**, 489–511 (1980).
5. Moore, B. W. *Biochem. biophys. Res. Commun.* **6**, 739–744 (1965).
6. Bock, E. *J. Neurochem.* **30**, 7–14 (1978).
7. Zomzely-Neurath, C. E. & Walker, W. A. in *Proteins of the Nervous System* (eds Bradshaw, R. A. & Schneider, D. M.) 1–57 (Raven, New York, 1980).
8. Marangos, P. J., Zomzely-Neurath, C., Luk, D. C. M. & York, C. J. *biol. Chem.* **250**, 1884–1891 (1975).
9. Tapia, et al. *Gut* **21**, 920 (1980).
10. Ferri, G.-L. et al. *Gut* **22**, 898 (1981).
11. Michetti, F., Miani, N., De Renzi, G., Caniglia, A. & Correr, S. *J. Neurochem.* **22**, 239–244 (1974).
12. Ludwin, S. K., Kosek, J. C. & Eng, L. F. *J. comp. Neurol.* **165**, 197–208 (1976).
13. Cocchia, D. *Cell Tissue Res.* **214**, 529–540 (1981).
14. Cocchia, D. & Miani, N. *J. Neurocytol.* **9**, 771–782 (1980).
15. Nakajima, T., Yamaguchi, H. & Takahashi, K. *Brain Res.* **191**, 523–531 (1980).
16. Cocchia, D. & Michetti, F. *Cell Tissue Res.* **215**, 103–112 (1981).
17. Cocchia, D., Michetti, F. & Donato R. *Nature* **294**, 85–87 (1981).
18. Stefansson, K., Wollmann, R. L., Moore, B. W. & Arnason, B. G. W. *Nature* **295**, 63–64 (1982).
19. Jessen, K. R. & Mirsky, R. *Nature* **286**, 736–737 (1980).
20. Smith, B. *The Neuropathology of the Alimentary Tract* (Arnold, London, 1972).
21. Bishop, A. E., Polak, J. M., Bloom, S. R. & Pearce, A. G. E. *J. Endocr.* **77**, 25–26 (1978).
22. Sternberger, L. *Immunocytochemistry*, 2nd edn (Wiley, New York, 1979).
23. Pickel, V. M., Jon, H. T. & Reis, D. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 659–663 (1975).

Hapten-specific T suppressor factor recognizes both hapten and I-J region products on haptenized spleen cells

M. Zembala*, G. L. Asherson & V. Colizzi*

Division of Immunological Medicine, Clinical Research Centre, Watford Road, Harrow HA1 3UJ, UK

The classical example of genetic restriction is associative recognition, the finding that cytotoxic T-cell killing requires matching between part of the major histocompatibility complex (MHC) of the stimulating cell and the target cell^{1,2}. Its analysis would be facilitated by the study of a system in which antigen-specific T suppressor factor (TSF) was genetically restricted in its interaction with haptenized cells. However, only one such system is known^{3,4}, although there are several systems in which the TSF must be syngeneic with other cells⁵⁻⁸. We now describe a genetically restricted TSF which arms an acceptor T cell. This acceptor cell shows some resemblance to the auxiliary T suppressor (Ts) cell, the acceptor T cell and the Ts, described elsewhere^{5,6,9}. In the presence of antigen (haptenized spleen cells) the armed acceptor cell releases a nonspecific inhibitor of the passive transfer of contact sensitivity (see Fig. 1). The important finding is that the antigen will only trigger the release of the nonspecific inhibitor when there is matching in the I-J region between the TSF and the haptenized cell used as a source of the triggering antigen.

In practice, the T acceptor (T_{acc}) cell is produced by immunization with picryl chloride (pic) or oxazolone. The contact sensitizer used is unimportant because it does not influence the specificity of the system (Table 1). Macrophages are removed by nylon wool filtration as they may act as acceptor cells¹⁰. The T_{acc} cell population is armed by incubation in anti-pic or anti-oxazolone TSF^{11,12}. The cells are then exposed to antigen (picrylated spleen cells) and the nonspecific inhibitor released into the supernatant is tested by its ability to block the transfer of contact sensitivity by immune cells incubated in it. This system is formally analogous to the reaginic system in which the roles of specific IgE, mast cell, soluble antigen and histamine are taken by specific TSF, the acceptor T-cell, haptenized spleen cells and the nonspecific inhibitor (see Fig. 1).

The first experiment investigated the main characteristics of the system. T acceptor cells were produced by immunization with picryl chloride or the unrelated contact sensitizer oxazolone, and were armed with anti-picryl TSF. After washing, picrylated spleen cells were added as a source of antigen and the supernatant was collected after 2 h. Table 1 (lines 3-5) shows that nonspecific inhibitor of the passive transfer of contact sensitivity was produced when the T_{acc} cells were generated by painting with picryl chloride or oxazolone, while normal spleen cells were ineffective, that is, immunization but not specific immunization was required for the generation of T_{acc} cells. Unseparated or nylon wool-purified T spleen cells could be used as a source of antigen (lines 3, 7) and irradiation at 2,000 rad did not affect the ability of picrylated spleen cells to trigger the release of nonspecific inhibitor.

We next investigated the requirement for genetic matching between the anti-picryl TSF and the picrylated cells used as a source of antigen. Immune T cells from CBA or BALB/c mice were used as an acceptor cell population and armed with anti-picryl TSF from CBA or BALB mice. Picrylated CBA or BALB spleen cells were then added as a source of antigen and the supernatant was collected at 2 h. Table 2 (line 3) shows that

Table 1 Acceptor cells require immunization but not specific immunization for their production

	TSF	Antigen	T _{acc}	Test cell		
1	Positive control			Ox CBA		77% ***
2	Negative control			Ox CBA		100% ***
3	Pic CBA	Pic CBA	Pic CBA	Ox CBA		20%
4	-	-	Ox CBA	Ox CBA		81% **
5	-	-	- CBA	Ox CBA		100% ***
6	-	-	Pic CBA	Ox CBA		98% ***
7	-	T _{nyl} pic CBA	-	Ox CBA		100% ***
8	-	-	-	Ox CBA		100% ***

Ear swelling (10⁻³ cm ± s.d.)

Anti-picryl TSF was produced by injecting picrylsulphonic acid (PSA) intravenously, painting with picryl chloride on day 6 and culturing lymphoid cells taken on day 7 for 48 h (ref. 12). Nylon wool-purified T spleen cells (T_{nyl}, 1.5 × 10⁸) taken from mice 4 days after painting with picryl chloride (Pic) or oxazolone (Ox) or from normal mice were used as acceptor cells. They were armed with anti-picryl TSF (4 ml) at 37 °C for 1 h, spun down, washed once and antigen added. The standard antigen was picrylated normal spleen cells (3 × 10⁷), prepared by lysis of red cells with Boyle's solution and picrylation with 10 mM PSA for 10 min (ref. 22). The armed acceptor cells and the picrylated spleen cells were centrifuged together and resuspended in 5 ml Eagle's minimal essential medium/Dulbecco's phosphate-buffered saline (MEM/PBS) with 2.5% inactivated foetal calf serum and incubated for 2 h with shaking every 40 min. Finally the nonspecific inhibitor in the supernatant was tested by incubating passive transfer cells (2.5 × 10⁸, 4 day, oxazolone immune lymph node and spleen cells) with nonspecific supernatant (5 ml) for 1 h, washing once and injecting 4.5 × 10⁷ into each of four to five CBA mice¹². The mice were immediately challenged with 1% oxazolone on both sides of both ears and contact sensitivity was assessed by the ear swelling at 24 h. The horizontal bars show the increase in ear thickness at 24 h in units of 10⁻³ cm ± s.d. The figures show the percent depression of contact sensitivity taking the positive control group incubated in medium as 0% and the negative control group which did not receive any cells as 100%. †The picrylated spleen cells used as antigen were irradiated at 2,000 rad from a cobalt source. The significance of the difference from the positive control by Student's *t*-test was **P* < 0.05, ***P* < 0.002 and ****P* < 0.001.

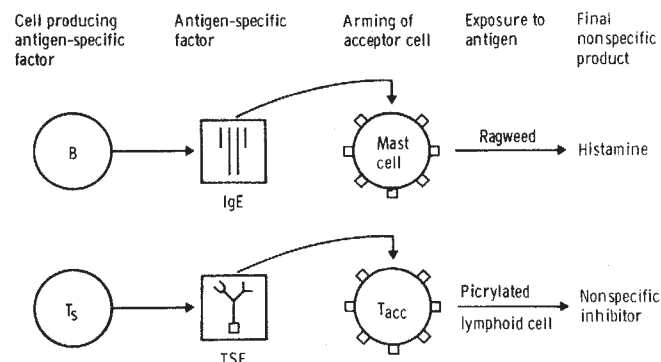


Fig. 1 Analogy between the IgE mast cell system and the T suppressor circuit involving Ts-eff, T suppressor factor and the T acceptor cell. In allergy to ragweed, the B cell makes specific IgE antibody. This is cytophilic and behaves like a mobile receptor which arms the mast cell. On exposure to ragweed the mast cell releases the final nonspecific products such as histamine. Similarly, in the T suppressor system to the picryl group, the T suppressor cell produces an anti-picryl T suppressor factor. This is cytophilic and arms the T acceptor cell. On exposure to antigen (picrylated lymphoid cells with the same I-J region as the cells producing the TSF) the T acceptor cell releases the final nonspecific products such as the inhibitor of the transfer of contact sensitivity. The diagram of the T suppressor factor illustrates its binding sites for the attachment to the T acceptor cell (□), the recognition of the picryl hapten (<) and the recognition of the product of the I-J locus (◊).

nonspecific inhibitor of the passive transfer of contact sensitivity was detected when the TSF, the acceptor T cell, the antigen (picrylated spleen cell) and the final test cell were all syngeneic. Lines 7, 9 and 11 of Table 2 show that syngeneity between the TSF and the antigen (picrylated spleen cell) was critical for triggering the release of the nonspecific inhibitor from the acceptor cell.

* Permanent addresses: Institute of Paediatrics, Copernicus School of Medicine, Wielicka 265, 30-663 Krakow, Poland (M.Z.); Institute of Microbiology, University of Pisa, 56100 Pisa, Italy (V.C.).

antigen was initially presented, and the antigen used to trigger the release of nonspecific inhibitor. An interesting elaboration is that the ability of antigen-specific TSF to give rise to anti-idiotypic T suppressor cells⁶ might be due to the *I-J* antigen of the TSF (or of the macrophage or T acceptor cell to which it attaches) acting as a signal for the induction of suppressor cells.

M.Z. and V.C. thank the Wellcome Trust and the Lega Nazionale contro i Tumori for financial support, Malcolm Brenner, Miroslav Malkovsky and Liz Simpson for valuable discussion and Valerie Stein and Madelaine Watkins for technical assistance.

Received 29 January; accepted 26 March 1982.

1. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 52-177 (1979).
2. Shearer, G. M., Remn, T. G. & Garbarino, C. A. *J. exp. Med.* **141**, 1348-1364 (1975).
3. Moorhead, J. W. *J. exp. Med.* **150**, 1432-1447, (1979).
4. Claman, N. H., Miller, S. D., Conlon, P. J. & Moorhead, J. W. *Adv. Immun.* **30**, 121-157, (1980).
5. Taniguchi, M. & Tokuhisa, T. *J. exp. Med.* **151**, 517-527 (1980).
6. Sunday, M. E., Benacerraf, B. & Dorf, M. E. *J. exp. Med.* **153**, 811-822 (1981).
7. Germain, R. N. & Benacerraf, B. *Springer Semin. Immunopath.* **3**, 93-128 (1980).
8. Germain, R. N. & Benacerraf, B. *Scand. J. Immun.* **13**, 1-10 (1981).
9. Sy, M. S., Miller, S. D., Moorhead, J. W. & Claman, H. N. *J. exp. Med.* **149**, 1197-1207 (1979).
10. Ptak, W., Zembala, M., Asherson, G. L. & Marcinkiewicz, J. *Int. Archs. Allergy appl. Immun.* **65**, 121-128 (1981).
11. Asherson, G. L., Zembala, M., Thomas, W. R. & Perera, M. A. C. C. *Immun. Rev.* **50**, 3-45 (1980).
12. Asherson, G. L., & Zembala, M. *Immunology* **42**, 1005-1013 (1980).
13. Asherson, G. L. & Zembala, M. *Ann. N. Y. Acad. Sci.* (in the press).
14. Feldmann, M. & Kontiainen, S. *Molec. cell. Biochem.* **30**, 177-193 (1980).
15. Murphy, D. B., Yamauchi, K., Habu, S., Eardley, D. D. & Gershon, R. K. *Immunogenetics* **13**, 205-213 (1981).
16. Okuda, K., Minami, M., Sher, D. H. & Dorf, M. E. *J. exp. Med.* **154**, 468-479 (1981).
17. Taniguchi, M., Takei, I. & Tada, T. *Nature* **283**, 227-228 (1980).
18. Noonan, F. P. & Halliday, W. J. *Cell. Immun.* **50**, 41-47 (1980).
19. Streilein, J. W. & Klein, J. *Proc. R. Soc. B* **207**, 461-474 (1980).
20. Bromberg, J. S., Benacerraf, B. & Greene, J. *J. exp. Med.* **153**, 437-449 (1981).
21. Czitrom, A. A., Sunshine, G. H. & Mitchison, N. A. *Immunogenetics* **11**, 97-102 (1980).
22. Mishell, B. B. & Shiigi, S. M. *Selected Methods in Cellular Immunology*, 98 (Freeman, San Francisco, 1980).

Epstein-Barr virus-specific cytotoxic T-cell clones restricted through a single HLA antigen

L. E. Wallace, A. B. Rickinson, M. Rowe & M. A. Epstein

Department of Pathology, University of Bristol Medical School, University Walk, Bristol BS8 1TD, UK

HLA-restriction of antiviral cytotoxic T-cell responses in man, precisely analogous to the H-2-restriction first documented in murine virus infections¹⁻³, has now been unequivocally demonstrated with influenza virus-specific⁴⁻⁷ and with Epstein-Barr (EB) virus-specific⁸⁻¹² cytotoxic T cells reactivated *in vitro*. If individual effector cells of this kind recognize viral antigen in association with one particular HLA determinant, then it follows that cytotoxic T-cell preparations from HLA-heterozygous donors must be polyclonal with discrete component populations each functionally linked to one of the relevant HLA-A or -B antigens. Clonal expansion of individual cells from such preparations in medium containing T-cell growth factor (TCGF-medium) not only offers a means of testing this hypothesis but should also provide monospecific populations ideal for the analysis of human cytotoxic T-cell function. We describe here the establishment *in vitro* of EB virus-specific human cytotoxic T-cell clones each recognizing the virus-induced lymphocyte-detected membrane antigen (LYDMA) in association with one particular HLA-A or -B determinant.

EB virus-specific cytotoxic T cells can be reactivated *in vitro* by co-culturing peripheral blood mononuclear cells from previously infected (seropositive) individuals with X-ray irradiated cells of the autologous EB virus-transformed B-cell line at particular responder: stimulator ratios^{13,14}. Such preparations are preferentially cytotoxic to the autologous stimulating line

and kill allogeneic EB virus-transformed cell lines in an HLA-A and -B antigen-restricted fashion⁸⁻¹⁴. Recent work in this laboratory has shown that cytotoxic T cells of this kind can be established as continuously growing TCGF-dependent cell lines without loss of antigen specificity or of HLA-restriction (L.E.W., M.R., A.B.R. and M.A.E., in preparation).

In the present experiments, EB virus-specific cytotoxic T cells prepared by *in vitro* reactivation from seropositive donor SW (HLA-A11, AW24, B7, BW35) were immediately transferred to TCGF-medium with an equal number of X-ray irradiated cells of the autologous EB virus-transformed cell line SW-EB, as feeders. Foci of proliferating cells appeared within 3 days and the resultant T-cell line, maintained thereafter by twice weekly subculture with feeder cells as above, showed a pattern of EB virus-specific HLA-A and -B antigen-restricted cytotoxicity identical to that of the original reactivated preparation. The results in Fig. 1 show that strong killing of the autologous cell line by TCGF-dependent effector cells was not accompanied by any significant reactivity against EB virus genome-negative human leukaemic cell lines HSB2 and K562, included as indicators of natural killer (NK)-like activity¹⁵. Furthermore, lysis of allogeneic EB virus-transformed cells was particularly strong when associated with either HLA-A11 (target CM) or BW35 (target EG) antigen matching and progressively weaker with matching through B7 (target JT) and through AW24 (target TH). This pattern of HLA antigen 'preferences' has been observed reproducibly in reactivated cytotoxic preparations from donor SW¹⁴.

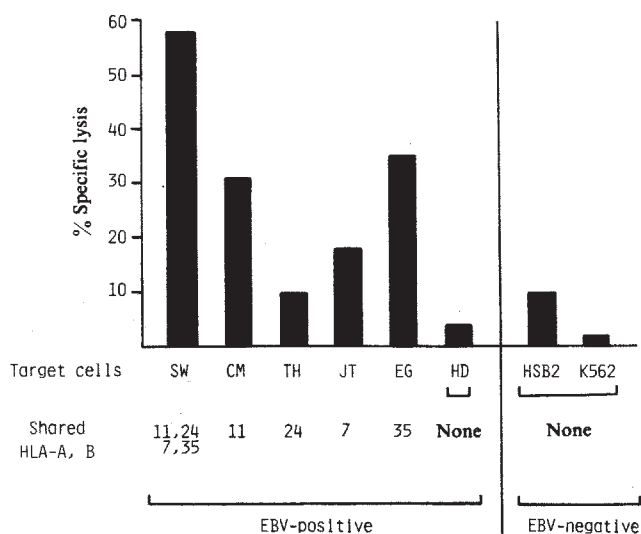


Fig. 1 EB virus-specific, HLA-A and -B antigen-restricted effector function of a TCGF-dependent cytotoxic T-cell line established from donor SW (A11, AW24; B7, BW35). Effector T cells were first obtained by culturing peripheral blood mononuclear cells for 10 days with X-ray irradiated cells of the autologous EB virus-transformed cell line (SW-EB) at a responder:stimulator ratio of 40:1 as described elsewhere^{13,14}. Thereafter the T cells were cultured in 2 ml Linbro wells at 2×10^5 cells per well together with 2×10^5 X-ray irradiated SW-EB cells per well in TCGF-medium (RPMI 1640 supplemented with 10% fetal calf serum and 25% TCGF prepared according to the method of Inoué *et al.*²⁴). Cells were collected every 3-4 days and recultured together with fresh stimulator cells in TCGF-medium as above. Thirteen days after their original exposure to TCGF, some of the cells were collected and incubated overnight in TCGF-free medium to remove any residual lectin. Their effector function was tested in a chromium release assay, against the autologous EB virus-transformed cell line, against a range of allogeneic EB virus-transformed cell lines whose degree of HLA-A and -B antigen-matching with the effector cells was as shown, and against EB virus genome-negative cell lines HSB2 and K562 included as indicators of NK-like activity. Results are expressed as percentage specific lysis of ⁵¹Cr-labelled target cells in a 5-h assay at an effector: target ratio of 10:1.

Table 1 Effect of the monoclonal antibodies W6/32 and Leu 2a on target cell lysis by clone G9

Monoclonal antibody	% Specific lysis of EB virus-transformed cells from donor:			
	SW (A11, AW24; B7, BW35)	BVR (A11, BW35)	EG (A2, A3; BW35, BW40)	SF (A2, A11; B7, BW44)
—	29	68	75	2
W6/32 (ref. 21)	2	15	23	13
Leu 2a (refs 22, 23)	0	15	16	6

Results of a chromium release assay, conducted as described in Fig. 1, using clone G9 effector cells and EB virus-transformed target cells from donors whose HLA-A and -B antigen type was as shown. Levels of percentage specific lysis seen in the standard assay are compared with those seen when target cells were first pretreated for 20 min at 37°C in the assay wells with a saturating concentration of W6/32 (1:200 final dilution of a purified immunoglobulin preparation) and with those seen when effector cells were first pretreated, again for 20 min at 37°C in the assay wells, with a saturating concentration of Leu 2a (1:200 final dilution of serum/ascites).

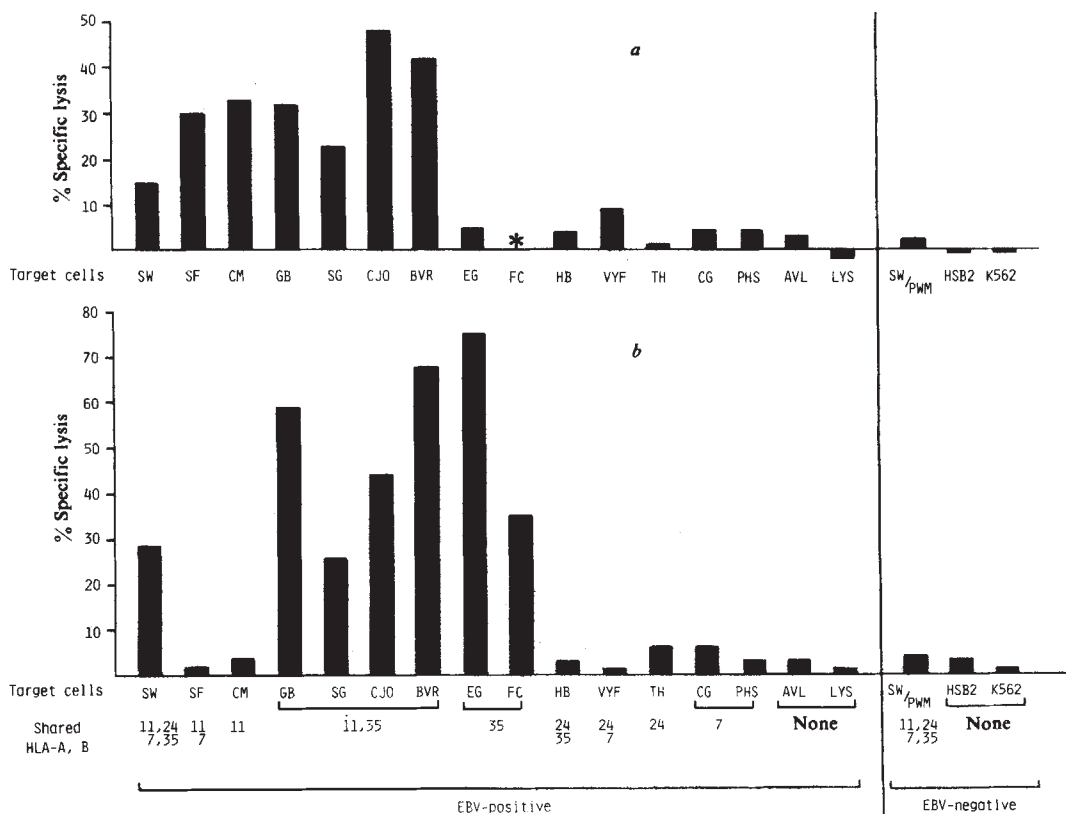
On day 12 following transfer of the effector cell preparation into TCGF-medium, a sample of the cells was cloned at 0.7 cells per well onto a feeder layer of X-ray irradiated SW-EB cells in U-shaped Microtest plate wells. The cultures were maintained with regular supplements of TCGF-medium and, after 2 weeks, those showing active growth were subcultured into replicate U-wells. Some of these replicate cultures provided effector cells for a preliminary screening assay of cytotoxicity against SW-EB and against HSB2 target cells. Of the 55 clones

screened in this way, 22 showed preferential lysis of the autologous cell line, and of these, 4 grew sufficiently well thereafter in TCGF-medium to permit extensive analysis of effector function and the establishment of cloned T-cell lines.

Results obtained with effector cells from clone C2 are illustrated in Fig. 2a, showing a representative range of targets only, and strongly suggest specificity for the viral antigen LYDMA in association with HLA-A11. In the tests overall, clone C2 killed 8/8 EB virus-transformed cell lines from donors typed as HLA-A11 but did not kill any of 15 corresponding lines from A11-negative donors even though many of these did possess AW24, B7 or BW35 in common with the effector cell donor SW. Moreover there was no lysis of EB virus genome-negative target cells, whether the NK-sensitive cell lines HSB2 and K562 or mitogen-stimulated lymphoblasts from autologous and from A11-matched donors were used. Two further clones, G13 and H8, gave results very similar to those obtained with clone C2. Another clone, G9, when tested against the same range of target cells, gave a different pattern of reactivity (Fig. 2b) suggesting specific recognition of LYDMA in association with HLA-BW35. Overall, clone G9 cells killed 9/10 EB virus-transformed target cell lines from donors typed as HLA-BW35 but none of the 11 corresponding lines tested from BW35-negative donors nor any of the relevant EB virus genome-negative target cells.

Previous studies have shown that EB virus-specific cytotoxic T-cell function can be blocked (1) by pretreatment of the target cells with saturating concentrations of monoclonal antibodies (such as W6/32) which bind to a framework determinant common to all HLA-A, -B and -C antigen molecules¹⁶, and (2) by pretreatment of the effector cells with the cytotoxic/suppressor T-cell-specific monoclonal antibody Leu 2a (L.E.W., M.R., A.B.R. and M.A.E., in preparation). The data in Table 1 clearly show that lysis of EB virus-transformed HLA-BW35-bearing

Fig. 2 EB virus-specific effector function of: a, clone C2, and b, clone G9 derived from the TCGF-dependent cytotoxic T-cell line of donor SW described in Fig. 1. Cloning was performed at an average of 0.7 cells per well seeded into U-shaped Microtest plate wells (Sterilin) containing 2.5×10^5 X-ray irradiated SW-EB cells in TCGF-medium. The cultures were re-fed with fresh TCGF-medium every 3–4 days and those showing active growth were first subcultured after 2 weeks and thereafter expanded into cell lines by regular subculture with X-ray irradiated SW-EB cells in fresh TCGF-medium. The cloned cell lines were tested in chromium release assays, exactly as described in Fig. 1 legend, for cytotoxicity against the autologous EB virus-transformed cell line, against a range of allogeneic EB virus-transformed cell lines whose degree of HLA-A and -B antigen-matching with the effector cells was as shown, against autologous pokeweed mitogen (PWM)-stimulated lymphoblasts and against EB virus genome-negative cell lines HSB2 and K562. Results shown for clone C2 are derived from three separate assays performed 35, 38 and 44 days after cloning; results shown for clone G9 are derived from three separate assays performed 31, 50 and 57 days after cloning. * Not tested.



target cells by clone G9 cells is subject to an exactly similar pattern of blocking by these monoclonal antibodies. This adds further support to the view that G9 is indeed a clone of HLA-restricted cytotoxic T cells. In this light, the inability of G9 cells to lyse the cell line from donor HB (HLA-A3, AW24, B14, BW35) is a most interesting result (Fig. 2b), particularly as this line did express all four relevant HLA antigens (as defined by serological typing) and was shown in other tests to be sensitive to LYDMA-specific cytotoxicity by HLA-B14-matched effector T cells (unpublished results). Whilst one could argue that preferential association of LYDMA with HLA-B14 (and possibly with A3 and AW24) on the HB target cell surface might explain the cells' insensitivity to BW35-restricted effector cells, it might alternatively be that donor HB possesses an unusual 'split' of BW35 which effector T cells can distinguish but which is not recognized by current typing sera. Studies with influenza virus-specific human cytotoxic T cells have recently provided examples of this phenomenon within the serologically defined antigens HLA-A2 and A3^{17,18}.

Clonal analysis of cytotoxic T-cell preparations *in vitro* was first achieved in the H-Y antigen-specific mouse model system¹⁹ and has more recently been investigated using this same antigen-specific system in man²⁰. The present report describes the first successful application of this technique in studying the human cytotoxic T-cell response to a natural pathogen. The results show that the response to EB virus is indeed polyclonal with individual reactivities recognizing the viral target antigen in the context of one particular HLA determinant. Since EB virus-specific effector T-cell preparations can be made from most previously infected individuals, it should be possible to establish a library of cytotoxic clones to cover the full range of HLA-A and -B antigen restrictions. Monospecific reagents of this kind could not only facilitate the biochemical analysis of cytotoxic T-cell function but also prove even more valuable as probes for HLA-A and -B antigen polymorphism.

We thank Elizabeth Gough, Ann Green and Jill Sweet for technical help and Dr B. A. Bradley and staff of the UK Transplant Services, SW Regional Transfusion Centre, Bristol, for HLA-A, -B and -C antigen typing of the cell donors. The gifts of monoclonal antibodies W6/32 and Leu 2a from Dr M. J. Crumpton, Imperial Cancer Research Fund Laboratories, and Dr R. B. Evans, Sloan-Kettering Institute for Cancer Research, respectively, are gratefully acknowledged. This work was supported by the MRC (grant G979/1123) and assisted by the Cancer Research Campaign, London, out of funds donated by the Bradbury Investment Corporation of Hong Kong.

Received 1 February; accepted 6 April 1982.

Product of a transferred *H-2L^d* gene acts as restriction element for LCMV-specific killer T cells

Anders Örn, Robert S. Goodenow & Leroy Hood

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

Peter R. Brayton*, Jerold G. Woodward†, Richard C. Harmon† & Jeffrey A. Frelinger†

Departments of *Neurology and †Microbiology, University of Southern California Medical School, Los Angeles, California 90033, USA

The major histocompatibility complex (MHC) of the mouse, *H-2* on chromosome 17, contains several families of genes encoding cell-surface molecules which have a role in mediating immune responses^{1,2}. The class I genes encode a family of homologous membrane proteins including the transplantation antigens K, D and L. These antigens reflect extensive genetic polymorphism which is apparent in the many different class I gene constellations or haplotypes found in mice (for example, BALB/c mice exhibit the *H-2^d* haplotype and their class I molecules are denoted K^d, D^d and L^d). Transplantation antigens serve as targets for T-cell killing in allogeneic immune responses such as *in vivo* graft rejection³ and destruction of allogeneic cells by cytotoxic T cells *in vitro*⁴. However, the physiological role of transplantation antigens may be to serve as restricting elements in virus-mediated T-cell killing of infected self cells. Virus infection of mice generates killer T cells whose receptors must interact with the foreign viral antigen and a class I molecule or restricting element for the cytotoxic effector function to be activated^{5,6}. Thus the T-cell receptor recognizes the viral antigen in the context of a class I molecule. To study the interaction between the T-cell receptor and the class I restricting element, we have used the mouse L-cell transformant 8-5 which expresses L^d molecules⁷ and the K7-65 transformant expressing K^d molecules (R.S.G. *et al.*, in preparation). Mouse L cells are fibroblasts derived from C3H mice of *H-2^k* haplotype and monoclonal antibodies can be used to distinguish *H-2^d* molecules from the endogenous *H-2^k* products. Recently, we have demonstrated that both L^d (ref. 8) and K^d molecules (unpublished data) expressed on transformed L cells can act as target antigens for alloreactive cytotoxic T lymphocytes. Here we show that the L^d molecule on transformed mouse L cells can serve as a restricting element in lymphocytic choriomeningitis virus (LCMV) infection, whereas its K^d counterpart cannot.

Recently, we described (ref. 7 and R.S.G. *et al.*, in preparation) the production of L-cell transformants expressing L^d and K^d molecules. Briefly, thymidine kinase-negative (*tk*⁻) L cells were co-transformed with the herpesvirus thymidine kinase gene alone or with the *tk* gene together with a cloned L^d or K^d gene. The transformed cells were selected for thymidine kinase (TK) activity and screened for the expression of L^d or K^d molecules by radioimmunoassay with monoclonal antibodies. The expression of L^d and K^d was confirmed by two-dimensional gel electrophoresis. In this manner we generated three types of *tk*⁺ target cell—those expressing K^k and D^k molecules (L-*tk*⁺ cells), those expressing K^k, D^k and K^d molecules (K7-65 cells) and those expressing K^k, D^k and L^d molecules (8-5 clone). Note that mice (or cells) of the *H-2^k* haplotype do not appear to express L antigens.

C3H and BALB/c mice were infected with LCMV to generate cytotoxic T cells, and spleen cells from immunized mice were assayed in a 4-h ⁵¹Cr-release assay against various target cells (Fig. 1). The cytotoxic T cells generated in C3H mice killed the LCMV-infected target cells from cloned 8-5 cells (K^kD^kL^d), L-*tk*⁺ cells (K^kD^k), and K7-65 cells (K^kD^kK^d) (Fig. 1a). The same cytotoxic T cells killed neither LCMV-infected

1. Zinkernagel, R. M. & Doherty, P. C. *Nature* **251**, 547–548 (1974).
2. Doherty, P. C. & Zinkernagel, R. M. *J. exp. Med.* **141**, 502–512 (1975).
3. Doherty, P. C., Blanden, R. V. & Zinkernagel, R. M. *Transplant Rev.* **29**, 89–124 (1976).
4. McMichael, A. J., Ting, A., Zweerink, H. J. & Askonas, B. A. *Nature* **270**, 524–526 (1977).
5. Biddison, W. E. & Shaw, S. *J. Immun.* **122**, 1705–1709 (1979).
6. Shaw, S. & Biddison, W. E. *J. exp. Med.* **149**, 565–575 (1979).
7. McMichael, A. J., Parham, P., Brodsky, F. M. & Pilch, J. R. *J. exp. Med.* **152**, 195s–203s (1980).
8. Rickinson, A. B., Wallace, L. E. & Epstein, M. A. *Nature* **283**, 865–867 (1980).
9. Misko, I. S., Moss, D. J. & Pope, J. H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4247–4250 (1980).
10. Moss, D. J., Wallace, L. E., Rickinson, A. B. & Epstein, M. A. *Eur. J. Immun.* **11**, 686–693 (1981).
11. Fukukawa, T. *et al. J. Immun.* **126**, 1697–1701 (1981).
12. Tsoukas, C. D. *et al. J. Immun.* **126**, 1742–1746 (1981).
13. Rickinson, A. B. *et al. Cancer Res.* **41**, 4216–4221 (1981).
14. Wallace, L. E. *et al. Cell. Immun.* **67**, 129–140 (1982).
15. Jondal, M. & Targan, S. *J. exp. Med.* **147**, 1621–1636 (1978).
16. Wallace, L. E., Moss, D. J., Rickinson, A. B., McMichael, A. J. & Epstein, M. A. *Eur. J. Immun.* **11**, 694–699 (1981).
17. Biddison, W. E., Ward, F. E., Shearer, G. M. & Shaw, S. *J. Immun.* **124**, 548–552 (1980).
18. Biddison, W. E., Shearer, G. M. & Shaw, S. *J. Immun.* **127**, 2231–2235 (1981).
19. von Boehmer, H. *et al. Eur. J. Immun.* **9**, 592–597 (1979).
20. Goulmy, E. *et al. J. exp. Med.* **152**, 182s–190s (1980).
21. Barnstable, C. J. *et al. Cell* **14**, 9–20 (1978).
22. Ledbetter, J. A. *et al. J. exp. Med.* **153**, 310–323 (1981).
23. Evans, R. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 544–548 (1981).
24. Inouye, H., Hank, J. A., Alter, B. J. & Bach, F. J. *Scand. J. Immun.* **12**, 149–154 (1980).

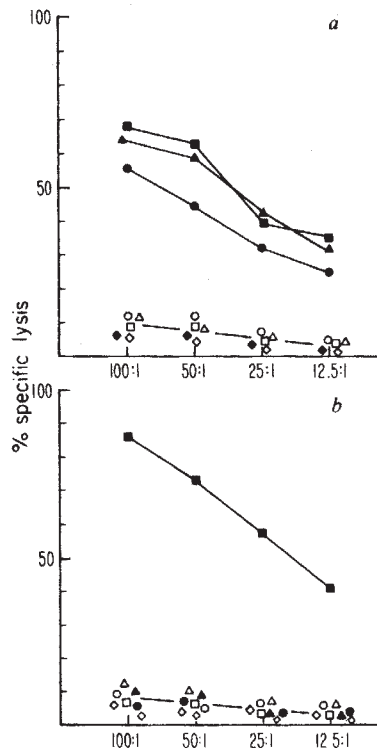


Fig. 1 Killing of LCMV-infected L-cell transformants by spleen cells from LCMV-infected C3H ($H-2^k$) and BALB/c ($H-2^d$) mice. Groups of two 6–8-week-old C3H/Bang and BALB/c mice were infected intraperitoneally with 2.5×10^2 plaque-forming units of the Armstrong strain of LCMV 7 days before cytotoxic assay. Spleens were removed and single cell suspensions prepared. L-cell transformants were grown in hypoxanthine-aminopterin-thymidine medium and Nulli-2A teratocarcinoma cells in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (FCS). The presence of L^d and K^d molecules on the transformants was monitored by radioimmunoassay using monoclonal antibodies as described previously⁷. Target cells were infected with LCMV 40 h before cytotoxicity assay at a multiplicity of infection of 0.5, then trypsinized, washed, and labelled for 1 h at 37 °C with 200 μ Ci 51 Cr in 0.5 ml RPMI supplemented with 10% FCS and 20 mM HEPES. After washing, 10^4 target cells were mixed in triplicates in microtitre plates with varying numbers of effector spleen cells in a total volume of 200 μ l. Plates were incubated at 37 °C, 5% CO_2 for 4 h, then centrifuged and 100 μ l of the supernatant were removed and counted in a γ -counter. Per cent specific lysis was calculated according to the formula: $\text{c.p.m.}_{\text{exp}} - \text{c.p.m.}_{\text{spont}} / \text{c.p.m.}_{\text{max}} - \text{c.p.m.}_{\text{spont}} \times 100$, where c.p.m. (counts per minute) (spontaneous) and c.p.m. (maximum) were determined in cultures containing medium or detergent, respectively; c.p.m. (experimental). The s.e.m. never exceeded 5%. The figure shows results for spleen effector cells from *a*, LCMV-infected C3H/Bang mice, and *b*, LCMV-infected BALB/c mice. Target cells: $\blacksquare$, 8-5 (L^d)-LCMV; $\square$, 8-5, uninfected; $\blacktriangle$, $L-tk^+$ -LCMV; $\triangle$, $L-tk^+$, uninfected; $\bullet$, K7-65 (K^d)-LCMV; $\circ$, K7-65, uninfected; $\blacklozenge$, Nulli-2A-LCMV; $\diamond$, Nulli-2A, uninfected.

Nulli-2A teratocarcinoma cells, which do not express class I molecules⁹, nor uninfected target cells. These data demonstrate that LCMV-specific T cells from C3H mice require the viral antigen and a K^k or D^k restricting element. In contrast, LCMV-restricted cytotoxic T cells from BALB/c mice killed LCMV-infected 8-5 cells but not any of the other target cells (Fig. 1*b*). These data demonstrate that BALB/c mice can use the L^d but not the K^d molecule as a T-cell restricting element in response to LCMV. This is consistent with reports on other $H-2$ -restricted systems which claim that L^d can act as a restricting element^{10–12}.

We then raised cytotoxic lymphocytes to LCMV in C- $H-2^{dm2}$ mice. These are mutant mice of the $H-2^d$ haplotype which fail to express an L^d molecule^{13,14}. Cytotoxic lymphocytes from the $dm2$ mice lysed neither LCMV-infected 8-5 ($K^kD^kL^d$) nor $L-tk^+$

(K^kD^k) cells (Fig. 2*a*). Control experiments with LCMV-induced cytotoxic T cells in C3H mice again demonstrated lysis of both the 8-5 and $L-tk^+$ cells as expected (Fig. 2*c*). A similar analysis with LCMV-specific killer T cells raised in B10.Q mice of the $H-2^q$ haplotype lysed neither the 8-5 nor $L-tk^+$ cells (Fig. 2*d*). This control experiment is interesting because mouse strains of the $H-2^q$ haplotype are known to express an L^q antigen serologically related to the L^d product^{15–17}. Thus, the T-cell receptor seems to interact with that portion of the L^d antigen which distinguishes it from its L^q counterpart.

We further tested the specificity of the L^d restricting element in LCMV-infected 8-5 cells by inhibiting the T-cell lysis of these cells with specific antibodies. Two L^d -specific alloantisera were raised by immunizing C3H mice with 8-5 cells as described in detail elsewhere¹⁸, and by immunizing C- $H-2^{dm2}$ mice with BALB/c spleen cells. In addition, we used two monoclonal anti- L^d antibodies (28-14-8S and 4-9.4), a monoclonal anti- D^d antibody (34-2-12), a monoclonal K^d antibody (6-27.10) and normal mouse serum. The results of these experiments are shown in Fig. 3*a, b*. All antibodies directed against the L^d molecule inhibited the LCMV-restricted cytotoxic T-cell killing of the L^d -transformed 8-5 cells. In contrast, normal mouse serum as well as monoclonal antibodies to the D^d and K^d antigens, did not inhibit this cytotoxic T-cell reaction. These data again suggest that the L^d molecule serves as a typical restricting element in LCMV-induced T-cell killing.

Thus, the gene transfer procedure results in the expression of class I molecules which are inserted in the membrane of a transformed cell in a normal manner so that they can act as restricting elements for virus-induced cytotoxic T cells. An analysis of various viral infections using the different transplantation antigens should allow us to delineate the roles of the class I molecules in specific virus infections. Moreover, this system, together with *in vitro* mutagenesis, should allow us to elucidate the nature of the interactions between the T-cell receptor and class I molecules. The major advantage of *in vitro* mutagenesis is the ability to generate mutants that carry alterations in specific regions of the class I molecules. In contrast, conventional mutants are selected only for altered T-cell recognition. Thus, our ability to transfer normal and altered class

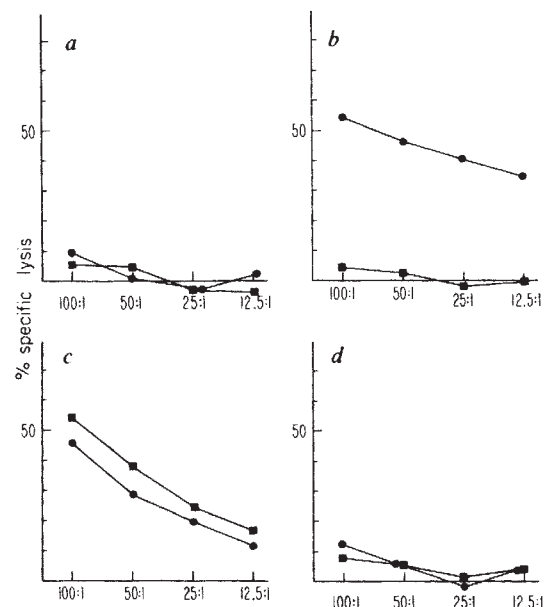


Fig. 2 Killing of LCMV-infected L-cell transformants by spleen cells from different LCMV-infected mouse strains. For experimental details see Fig. 1 legend. Spleen effector cells from LCMV-infected mice: *a*, C- $H-2^{dm2}$ ($H-2^d$ mutant); *b*, BALB/c ($H-2^d$); *c*, C3H/Bang ($H-2^k$); *d*, B10.Q ($H-2^q$). Target cells: $\bullet$, 8-5 (L^d)-LCMV; $\blacksquare$, $L-tk^+$ -LCMV. Uninfected target cells never showed >10% lysis.

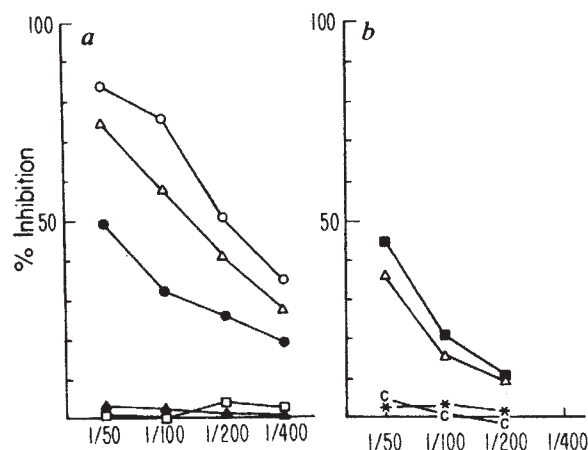


Fig. 3 Inhibition of *H*-2-restricted LCMV-specific cytotoxicity by antibodies directed against different *H*-2 antigens. Various antisera or monoclonal reagents were added at the start of the Cr-release assay in the dilutions shown. The cytotoxic reaction inhibited was that of LCMV-infected BALB/c spleen cells against LCMV-infected 8-5 cells. Per cent inhibition was calculated as follows:

$$1 - \frac{\% \text{ specific lysis in the presence of antibody}}{\% \text{ specific lysis in the absence of antibody}} \times 100$$

a and *b* show two separate experiments: *a*, 30.5% lysis in the absence of blocking antibody at effector/target ratio of 50:1. Blocking antibodies: ○, C3H/Bang anti-8-5 alloantisera; △, 28-14-8S (anti-*L*^d monoclonal antibody); ●, 4-9.4 (anti-*L*^d monoclonal antibody); ▲, 34-2-12 (anti-*D*^d monoclonal antibody); and □, normal mouse serum. The specificities of these monoclonal antibodies are described elsewhere^{8,19}. *b*, 41.0% lysis in the absence of blocking antibody at effector/target ratio of 100:1. Blocking antibodies: ■, C-*H*-2^{dm2} anti-2BALB/c alloantisera; △, 28-14-8S (anti-*L*^d monoclonal antibody); C, 6-27.10 (anti-*K*^d monoclonal antibody); and *, normal mouse serum.

I genes into foreign cells provides us with one of the first opportunities to elucidate the molecular basis of cell-cell interactions.

This work was supported by NIH grants GM 06965, CA 22662, CA 26199 and CA 25911. We thank Drs David Sachs and Ted Hansen for monoclonal antibodies and antisera; Dr Roger Melvold for supplying the C-*H*-2^{dm2} mice and Dr Peter Stern for the Nulli-2A cell line. R.S.G. is a Senior Lievre Fellow of the California Division of the American Cancer Society. J.G.W. is a recipient of a National Arthritis Foundation Post-doctoral Fellowship. R.C.H. and P.R.B. are supported by NRSA-NS-07149 and J.A.F. is the recipient of an American Cancer Society Faculty Research Award.

Received 23 March; accepted 14 April 1982.

- Klein, J. *Science* **203**, 516-521 (1979).
- Snell, G. D. *Science* **213**, 172-178 (1981).
- Snell, G. D., Dausset, J. & Nathenson, S. (eds) *Histocompatibility* (Academic, New York, 1976).
- Alter, B. J. *et al.* *J. exp. Med.* **137**, 1303-1309 (1973).
- Doherty, P. C. & Zinkernagel, R. M. *J. exp. Med.* **141**, 502-507 (1975).
- Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51-177 (1979).
- Goodenow, R. S. *et al.* *Science* **215**, 677-679 (1982).
- Woodward, J. G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Stern, P., Martin, G. R. & Evans, M. J. *Cell* **6**, 455-465 (1975).
- Biddison, W. E., Hansen, T. H., Levy, R. B. & Doherty, P. C. *J. exp. Med.* **148**, 1678-1686 (1978).
- Blanden, R. V., McKenzie, I. F. C., Kees, U., Melvold, R. W. & Kohn, H. I. *J. exp. Med.* **146**, 869-880 (1977).
- Levy, R. B., Shearer, G. M. & Hansen, T. H. *J. Immun.* **121**, 2263-2269 (1978).
- McKenzie, I. F. C., Morgan, G. M., Melvold, R. W. & Kohn, H. I. *Immunogenetics* **4**, 333-347 (1977).
- Hansen, T. H. *et al.* *J. Immun.* **126**, 1713-1716 (1981).
- Hansen, T. H. & Sachs, D. H. *J. Immun.* **121**, 1469-1472 (1978).
- Huang, C.-M., Huang, S. & Klein, J. *Immunogenetics* **9**, 173-182 (1979).
- Ozato, K., Hansen, T. H. & Sachs, D. H. *J. Immun.* **125**, 2473-2477 (1980).
- Woodward, J. G. *et al.* in *B and T Cell Tumors: Biological and Clinical Aspects* (eds Vitetta, E. & Fox, C. F.) (Academic, New York, in the press).
- Ozato, K., Meyer, R. & Sachs, D. H. *Transplantation* (in the press).

Sensitivity and resistance of human tumour cells to interferon and $rI_n \cdot rC_n$

Stanley L. Lin*, James J. Greene & P. O. P. Ts'o

The Johns Hopkins University, School of Hygiene and Public Health, Division of Biophysics, 615 North Wolfe Street, Baltimore, Maryland 21205, USA

W. A. Carter

Herbert L. Orlovitz Institute, Hahnemann Medical College and Hospital, Philadelphia, Pennsylvania 19102, USA

Interferon exerts, in addition to its well recognized antiviral activity, an antiproliferative effect^{1,2} and has been found to reduce plasminogen activator secretion³, to re-establish contact inhibition⁴ and to inhibit the growth in semi-solid agar⁵ of neoplastic cells. Double-stranded RNA (dsRNA) is both an interferon inducer and an activator of two interferon-induced enzymes, a protein kinase and a 2',5' oligoadenylate synthetase⁶. However, the molecular bases for the antiproliferative and antineoplastic effects of interferon, and their relationship to dsRNA, are unknown. We have now characterized the antiproliferative effects of interferon and dsRNA on the human HT1080 fibrosarcoma⁷ and RT4 epidermal carcinoma⁸ cell lines. Luria-Delbrück fluctuation analysis⁹ of interferon resistance indicates that the resistance trait is a non-randomly acquired phenotype which was probably induced by interferon at the time of selection. The interferon-resistant (IFN^r) cells retained sensitivity to the antiviral effects of interferon, as well as to the antiproliferative effects of dsRNA.

Clonal populations of interferon-sensitive (IFN^s) cells were isolated from both HT1080 and RT4 cell lines. The HT1080 (HT1080-2) and RT4 (RT4-1) clonal lines demonstrated a complete growth inhibition at 400 and 80 IU ml⁻¹ human fibroblast interferon (HuIFN-β) respectively, when examined over a 72 h period (data not shown). The fluctuation analysis was initiated by establishing 10 independent cultures from each original colony and continuously growing each culture for 16 population doublings, as described in Table 1 legend, before plating in selective medium. Results are shown in Table 1. In all cultures of the HT1080-2 population, 3.7-4.5% of the cells were capable of colony formation in interferon-containing medium (1,000 IU ml⁻¹). There was no difference in average colony size or cell morphology between control and interferon-treated cells. The frequency distribution of the number of resistant colonies obtained from this series of sister cultures exhibited a variance of 7.73. The variance of the distribution of the number of colonies in a large single culture (v_2 , random sampling control) was 2.04. After correction for the random sampling control, the variance-to-mean ratio [v.m.r., ($v_1 - v_2$)/mean] of the frequency distribution of the number of resistant colonies obtained from the series of cultures was 0.15. In contrast, when RT4-1 cells were similarly treated, interferon-resistant colonies arose at a frequency of 0.11%, 30-40-fold less than the frequency observed for HT1080-2 cells (Table 1). However, the v.m.r. of the distribution obtained from the series of 10 sister cultures was low—0.31. Simultaneous examination of the HT1080-2 cultures for acquisition of 6-thioguanine resistance (6TG^r), an acquired trait of known spontaneous origin, resulted in a corrected v.m.r. of 6.79 (Table 1). The high v.m.r. is indicative of a randomly acquired trait and indicates that 16 population doublings were sufficient to distinguish between randomly and non-randomly acquired phenotypes. Using the p_0 method of Luria and Delbrück⁹, a mutation rate of 1.8×10^{-8} mutants per cell per generation can be calculated from the values presented in Table 1 for the acquisition of 6TG^r. This rate is consistent with known mutation rates obtained for various loci in other systems¹⁰, and thus verifies our procedure.

* Present address: The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA.

Examination of isolated colonies derived from resistant populations revealed that IFN^r cells possessed the same population doubling times in the presence and absence of interferon at concentrations normally resulting in complete cessation of growth (Fig. 1a, c). However, consistent with the notion that IFN^r is an induced trait, resistant colonies reverted to a sensitive state, albeit after varying lengths of time. One clone of HT1080-IFN^r cells retained the resistance trait after 4 weeks growth in interferon-free medium and reverted to a sensitive state by 15 weeks after removal of interferon from the culture medium (Fig. 1a), whereas a clone of resistant RT4 cells was fully resistant to interferon immediately after isolation, but reverted completely by 4 weeks after removal of interferon from the culture medium (Fig. 1c). When used as assay cells to titrate an identical solution of interferon by the dye uptake method of Finter¹¹, resistant cells during their resistant state were at least as sensitive as their parental populations to the antiviral effects of interferon when vesicular stomatitis virus was used as the challenge virus (Table 2), suggesting that resistance to the antiproliferative effects of interferon do not necessarily correlate with changes in sensitivity to its antiviral effects. It should be mentioned, however, that different viruses can be affected by different mechanisms pertaining to the antiviral state¹²⁻¹⁴.

In view of the participation of dsRNAs in the establishment of the antiviral state and the toxicity of exogenous dsRNA¹⁵, we characterized the sensitivities of the interferon-resistant lines towards the antiproliferative effects of the synthetic dsRNA, rI_n·rC_n (Fig. 1). Resistant lines of both HT1080-2 (Fig. 1b) and RT4-1 (Fig. 1d) displayed the same sensitivity to

Table 1 Luria-Delbrück fluctuation analysis for the acquisition of interferon resistance

Culture	HT1080		RT4
	IFN ^r colonies per 10 ³ cells	6TG ^r colonies per 10 ⁶ cells	IFN ^r colonies per 10 ³ cells
1	39.3	7.6	126
2	36.9	2.4	145
3	38.2	6.1	127
4	35.7	14.2	127
5	41.1	0.0	97
6	42.5	0.0	102
7	44.6	18.0	96
8	39.1	0.0	119
9	37.1	2.5	122
10	35.9	2.7	108
Variance	7.73	35.25	224.09
Mean	39.04	5.35	116.90
Random sampling variance	2.04	0.16	187.40
Corrected variance/mean	0.15	6.56	0.31

HT1080 cells were cloned in semi-solid agar as described by Laug *et al.*²¹. A vigorously growing colony was isolated and transferred to a 35 mm dish containing 2 ml growth medium, disburbed by repeated pipetting and allowed to attach and grow. The resultant monolayer was detached from the plate by trypsinization and was transferred to a 75 cm² flask for serial propagation. Cells comprising the resultant confluent monolayer were detached by trypsinization and 10 independent cultures were initiated by inoculating 10 35 mm dishes at 100 cells per dish. On reaching confluence, the total number of cells from each culture was detached by trypsinization and transferred to a separate 75 cm² flask to maintain the 10 cultures independently. After a total of 16 population doublings from initial establishment of the independent cultures, confluent monolayers in 75 cm² flasks were detached by trypsinization and the frequency of IFN^r cells in each independent culture determined by placing cells in liquid medium containing HuIFN-β at a plating density of 1 × 10⁴ cells per 100 mm dish. Fluctuation analysis for the acquisition of 6TG^r was simultaneously performed using selective conditions described previously²². An RT4 clone was isolated by the agar overlay method described in ref. 23, transferred to a 35 mm dish containing 2 ml growth medium and then handled in identical fashion to the HT1080 clone. RT4 cells were plated at a density of 1 × 10⁵ cells per 100 mm dish in interferon-containing medium. Selective medium consisted of growth medium containing HuIFN-β (1,000 IU ml⁻¹) at a plating density of 1 × 10⁴ cells per 100 mm dish. Fluctuation analysis or 6TG (3 μg ml⁻¹). Results shown are corrected for plating efficiency. Random sampling variance was determined by analysing the frequency distribution observed in a single large culture in the same way as for independent cultures. Correction for sampling error was performed by subtracting the random sampling variance from the variance of the independent cultures before division by the mean. IFN-β (fibroblast-derived) used in these studies had specific activities of 2 × 10⁶ (HEM Research, Rockville, Maryland) to 2 × 10⁷ (Dr W. A. Carter) IU per mg protein after purification by sequential concanavalin A-Sepharose → phenyl-Sepharose affinity chromatographies²⁴.

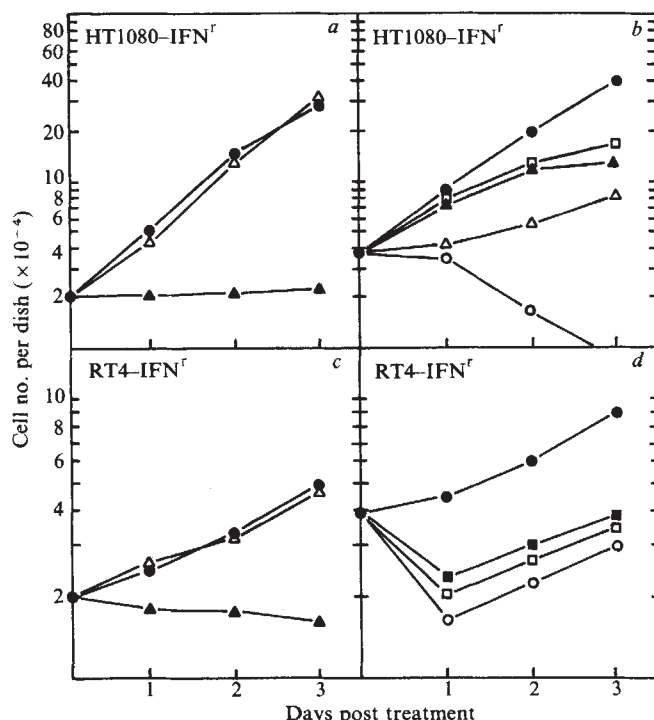


Fig. 1 Heritability of resistance to the growth inhibitory effects of interferon, and effects of rI_n·rC_n on the growth of IFN^r cells in culture. Colonies growing in interferon-containing medium were isolated and grown in the presence of medium containing 1,000 (HT1080) or 100 (RT4) IU ml⁻¹ HuIFN-β. After ~16 population doublings, HuIFN-β was removed from the culture medium and cells were examined for interferon resistance at various passages after growth in interferon-free medium (a, c) or for sensitivity to rI_n·rC_n-mediated toxicity. (a) Growth of HT1080-IFN^r cells in the absence (●) or presence of HuIFN-β (1,000 IU ml⁻¹) after culture for 4 (Δ) and 15 weeks (▲) in interferon-free medium; (c), growth of RT4-IFN^r cells in the absence (●) or presence of HuIFN-β (100 IU ml⁻¹) after culture for 0 (Δ) and 4 (▲) weeks in interferon-free medium. For measuring dsRNA-mediated toxicity, 5 × 10⁴ (HT1080-IFN^r) or 1 × 10⁵ (RT4-IFN^r) cells were plated out per 35 mm dish and allowed to attach overnight. The following day (time 0), cell number per dish was determined, remaining monolayers were refed with various concentrations of rI_n·rC_n and cell number per dish was determined every 24 h thereafter. (b), HT1080-IFN^r cells grown in 0 (●), 0.1 (□), 0.2 (▲), 0.5 (Δ) and 2.5 (○) mM rI_n·rC_n; (d), RT4-IFN^r cells grown in 0 (●), 0.1 (■), 5.0 (□) and 10.0 (○) μM rI_n·rC_n.

rI_n·rC_n-mediated toxicity as did their respective parental HT1080-2 and RT4-1 lines (not shown). For HT1080-IFN^r populations, 0.5–1.0 mM rI_n·rC_n was required to inhibit cell proliferation completely, with rI_n·rC_n concentrations >1 mM producing marked cell death within 24 h. In these conditions, interferon concentrations secreted in the medium did not exceed 3 IU ml⁻¹. Detectable rI_n·rC_n-induced toxicity in RT4-IFN^r cells was observed at concentrations as low as 0.1 μM, with no induced interferon being detectable in the medium. However, in contrast to the prolonged cytotoxic effect observed in HT1080-2 cells, some clones of RT4-IFN^r cells could recover from dsRNA-mediated toxicity after 24 h and exhibit the same growth rate as before treatment. This recovery was not the result of rapid degradation of rI_n·rC_n as re-addition of rI_n·rC_n to the culture medium at 24 or 48 h did not alter the recovered growth rate (data not shown). Addition of antisera, directed against HuIFN-β (final concentration 100 neutralizing units per ml), to the culture medium of rI_n·rC_n-treated cells did not reduce or otherwise alter the extent of cytotoxicity induced by rI_n·rC_n (not shown), suggesting that low interferon concentration was not involved in rI_n·rC_n-mediated toxicity.

The Luria-Delbrück fluctuation analysis, originally developed for microbial systems, provides a statistical method whereby the acquisition of a certain phenotype by a population

Table 2 Sensitivity of IFN^s and IFN^r lines to the antiviral effects of interferon

Cell line	Phenotype	Titre
HT1080-2	IFN ^s	1.5
HT1080-2	IFN ^r	4.5
RT4-1	IFN ^s	30.2
RT4-4	IFN ^r	28.0

Confluent monolayers (1.5 cm diameter) of IFN^s or IFN^r cells from both HT1080 and RT4 lines were used as assay cells for the titration of a stock solution of interferon (concentration 50 IU ml⁻¹), using the dye uptake method of Finter¹¹ with vesicular stomatitis virus as the challenge virus.

of cells can be determined to be random (spontaneous) or non-random (induced)⁹. When a series of separate cultures, all initiated from a single clone, is allowed to undergo a suitable number of population doublings, the variation of the frequency of resistant colonies obtained in each culture will be large if the resistant trait is randomly acquired, or small if the trait is non-randomly acquired. Theoretically, the v.m.r. will be ≤ 1 if the trait is acquired non-randomly, or $\gg 1$ if the trait is acquired randomly. The acquisition of interferon resistance when analysed in this way, provided v.m.r.s of < 1 , suggesting that interferon resistance does not result from a spontaneous event, such as mutation or chromosomal rearrangement, but rather is a phenotype induced by interferon treatment. This conclusion is supported by the demonstration reported here, and elsewhere¹⁶, that resistant cells, previously sensitive to the growth inhibitory effects of interferon, revert to a sensitive state. As no difference in average size was observed for colonies formed in control and interferon-containing medium, induction of resistance must occur rapidly, within 24–48 h after interferon exposure. Alternatively, cells in the population might exist in either an IFN^s or IFN^r state, depending on their physiological or developmental status. Addition of interferon would both select for the subpopulation of pre-existing resistant cells and maintain the resistant trait.

Our results contrast with those of Gresser *et al.*¹⁷, who by similar analysis concluded that the resistance phenotype was randomly acquired and appeared at a frequency consistent with somatic mutation. However, Gresser *et al.* obtained their high v.m.r. from comparison of resistance displayed by different clones isolated from a mass tumour cell culture. Recent observations that clonal isolates from a population of tumour cells can vary widely in sensitivity to the antiproliferative effects of interferon¹⁶ indicate that such interclonal comparison cannot be used for determining the randomness of acquisition of interferon resistance. Moreover, the resistant cells we used, as opposed to those described by Gresser *et al.*¹⁷, retained their sensitivity to the antiviral effects of interferon, indicating that resistance was not due simply to the absence of an interferon receptor, and is consistent with previous reports where the antiviral state could be induced by interferon in Rous sarcoma virus-transformed human fibroblasts resistant to the growth inhibitory effects of interferon¹⁸.

Resistant cells retained a sensitivity towards dsRNA-induced toxicity that did not differ from that exhibited by the IFN^s parental cell lines. This is an important consideration in the development of chemotherapeutic regimens for the treatment of cancer with interferon and it suggests that interferon used in combination with dsRNA would be much more effective in the treatment of cancer than interferon alone. In addition, the data argue against a role for the known interferon-induced, dsRNA-dependent enzymes in the antiproliferative effects of interferon. No reduction in interferon-mediated growth inhibition was observed in mitogen-stimulated normal human fibroblasts pretreated with interferon for periods of up to 72 h (ref. 19) and continuous subculture of a sensitive line of normal human fibroblasts in the presence of HuIFN- β for up to 6 weeks did not result in the appearance of a resistant population²⁰, suggesting that normal cells do not readily acquire inter-

feron resistance. This difference between the ability of normal and neoplastic cells to acquire resistance to the growth inhibitory effects of interferon might be related to the mechanisms of uncontrolled growth characteristic of tumour cells. Further delineation of interferon-mediated growth inhibition could be important in elucidating a fundamental difference between normal and neoplastic cells, as well as in establishing the molecular basis for the antiproliferative effects of interferon.

Received 11 January; accepted 5 April 1982.

- Stewart, W. E. II, Gresser, I., Tovey, M., Bandu, M. T. & Le Goff, S. *Nature* **262**, 300–302 (1976).
- Pfeffer, L. M., Murphy, J. S. & Tamm, I. *Expl Cell Res.* **121**, 111–120 (1979).
- Schroder, E., Chou, I.-N., Jaken, S. & Black, P. H. *Nature* **276**, 828–829 (1978).
- Knight, E. Jr *J. Cell Biol.* **56**, 846–849 (1973).
- Gresser, I., Thomas, M.-T. & Brouty-Boyd, D. *Nature* **231**, 20–21 (1971).
- Farrel, P. J. *et al. Proc. natn. Acad. Sci.* **75**, 5893–5897 (1978).
- Rasheed, S., Nelson-Rees, W. A., Toth, E. M., Arnstein, P. & Gardner, M. B. *Cancer* **33**, 1027–1033 (1974).
- Rigby, C. C. & Franks, L. M. *Br. J. Cancer* **24**, 746–754 (1970).
- Luria, S. E. & Delbrück, M. *Genetics* **28**, 491–511 (1943).
- Barrett, J. C., Crawford, B. D. & Ts'o, P. O. P. in *Advances in Modern Environmental Toxicology*, Vol. 1 (eds Mishra, N., Dunkel, V. & Mehlman, M.) 467–501 (Senate, New Jersey, 1980).
- Finter, N. J. *gen. Virol.* **5**, 419–427 (1969).
- Nilsen, T. W., Wood, D. L. & Baglioni, C. *Nature* **286**, 178–180 (1980).
- Cearnicki, C. W., Sreevalsan, T., Friedman, R. & Panet, A. J. *Virol.* **37**, 827–831 (1981).
- Ankel, H., Krishnamurti, C., Besancon, F., Stefanos, S. & Falcoff, E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2528–2532 (1980).
- Stewart, W. E. II, de Clerq, E., Billiau, A., Desmeyer, J. & de Somer, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1851–1854 (1972).
- Creasey, A. A., Bartholomew, J. C. & Merigan, T. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1471–1475 (1980).
- Gresser, I., Bandu, M. -T. & Brouty-Boyd, D. *J. natn. Cancer Inst.* **52**, 553–559 (1974).
- Kuwata, T., Fuse, A. & Morinaga, N. *J. gen. Virol.* **33**, 7–15 (1976).
- Lin, S. L., Ts'o, P. O. P. & Hollenberg, M. D. *Biochem. biophys. Res. Commun.* **96**, 168–174 (1980).
- Lin, S. L. thesis, Johns Hopkins Univ. (1982).
- Laug, W. E., Jones, P. A. & Benedict, W. F. J. *natn. Cancer Inst.* **54**, 173–179 (1975).
- Moray, D. W. & Ts'o, P. O. P. *Mutat. Res.* **70**, 221–229 (1980).
- Thompson, L. H. & Baker, R. M. *Meth. Cell Biol.* **6**, 209–281 (1973).
- Davey, M. W., Sulkowski, E. & Carter, W. A. *J. biol. Chem.* **251**, 7620–7625 (1976).

Stimulation of adrenal mitogenesis by N-terminal proopiomelanocortin peptides

F. E. Estivariz & F. Iturriza

Centro de Estudios Endocrinos, Facultad de Ciencias Medicas, Universidad Nacional de la Plata, Casilla de Correo 455, 1900 La Plata, Argentina

C. McLean, J. Hope & P. J. Lowry

Pituitary Hormone Laboratory, Department of Chemical Pathology, St Bartholomew's Hospital, London EC1 7BE UK

Although it was generally believed that corticotropin (ACTH) maintained both the size of the adrenal gland and its level of steroid production¹, there is a growing body of evidence for a specific factor(s) distinct from ACTH which may be responsible for the stimulation of adrenocortical growth and proliferation. While direct neural influences have been proposed to be responsible for the compensatory adrenal growth following unilateral adrenalectomy², several observations have suggested the involvement of peptides such as angiotensin³ and vasopressin⁴, which have other hormonal effects, or, for example, fibroblast growth factor⁵, which has more general mitogenic actions. Moreover ACTH inhibits cell proliferation *in vitro*^{6–8}, and physiological doses *in vivo* cannot induce the compensatory growth and hyperplasia seen in the remaining adrenal gland after unilateral adrenalectomy⁹. These observations coupled with the lack of effect of chronic treatment with an ACTH antiserum on adrenal size¹⁰ and the existence of adrenal weight-maintaining activity in the plasma of patients with ACTH-secreting adenomas¹¹, have led to the search for a pituitary factor related to ACTH which is capable of stimulating adrenal growth. We now report that peptides derived from the N-terminal non- γ -melanocyte-stimulating hormone (MSH) portion of the precursor of ACTH, proopiomelanocortin, are potent stimulators of adrenal DNA synthesis *in vitro* and mitosis *in vivo* and may thus be involved in the physiological control of adrenal growth.

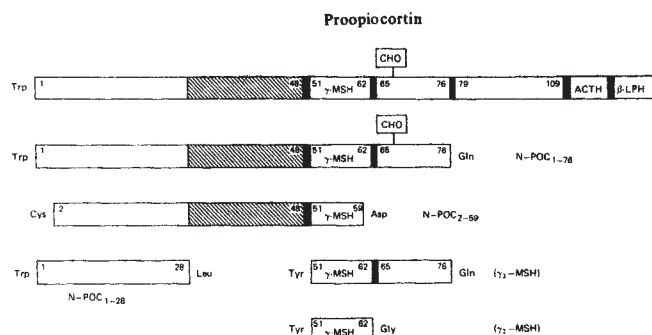


Fig. 1 Diagram of the human N-POC peptides used in this study in relation to each other and to their precursor proopiomelanocortin. The numbering has been assigned from chemical characterization of isolated peptides¹⁸ and cDNA¹⁹ sequence data. The γ_2 - and γ_3 -MSH synthetic peptides were a gift from Dr Nicholas Ling. The hatched area corresponds to the region recognized by the anti-serum¹⁷ used in this study and the boxed CHO, the polysaccharide moiety.

ACTH is derived from a precursor which contains lipotropin (LPH)/endorphin as the C-terminal portion and an N-terminal glycopeptide (N-POC) of some 100 amino acids (Fig. 1) which contains a third melanotropin sequence (γ -MSH)¹². A human pituitary glycopeptide (N-POC₁₋₇₆) comprising almost all of the non-ACTH non-LPH/endorphin region, which is inactive in maintaining adrenal weight and possesses weak melanotropic activity¹³, has recently been found to potentiate the ACTH-induced secretion of glucocorticoids and mineralocorticoids from the adrenal cortex¹⁴ by increasing RNA synthesis¹⁵. The potentiation of steroidogenesis can also be demonstrated with small synthetic peptides containing the γ -MSH sequence found in the C-terminal part of the glycopeptide¹⁶. Initial attempts to demonstrate adrenal weight maintenance with human N-POC₁₋₇₆ given either alone or with physiological doses of ACTH to hypophysectomized rats were unsuccessful¹³.

The generation of an antiserum directed against the middle portion of this peptide (residues 29–48)¹⁷ allowed us to test its effects on the adrenal in intact rats. Rats were injected daily for 7 days with the anti-N-POC antiserum, an antiserum which cross-reacts fully with both ACTH and β -LPH, or normal rabbit serum. Thymidine incorporation was also measured in these experiments as a more specific index of DNA synthesis. A decrease in adrenal weight (14%) was observed in the anti-N-

POC antiserum-treated group although the result was not significant (results not shown). However, as shown in Fig. 2a, there was a significant decrease in thymidine incorporation in the anti-N-POC antiserum-treated group (B), whereas the anti-ACTH/LPH antiserum-treated group (C) did not differ significantly from controls (A). The efficacy of the anti-ACTH/LPH antiserum treatment was shown by the significantly lower levels of plasma corticosterone in this group.

In preliminary experiments, however, we had also failed to observe any effects of several days' treatment with N-POC₁₋₇₆ in rats *in vivo* using thymidine incorporation as the measure of DNA synthesis. This led us to suspect that N-POC₁₋₇₆ could be a precursor for the mitogenic factor—this indeed appeared to be the case. Figure 2b confirms the absence of an effect after seven daily injections of the intact N-POC₁₋₇₆ (E) but pretreatment of the peptide with trypsin revealed significant activity (F). In this experiment, dexamethasone blockade was used in order to preserve as fully as possible the endocrine status of the animals whilst avoiding the stress-induced release of ACTH-related peptides due to the daily injections. Recently¹⁸ we have purified and chemically characterized smaller fragments from the N-terminus of N-POC₁₋₇₆, including N-POC₁₋₂₈ and N-POC₂₋₅₉, and have compared the effects of these peptides with synthetic γ_2 -MSH, γ_3 -MSH and ACTH on isolated adrenal cells. Whereas γ_2 and γ_3 -MSH had no effect in this system

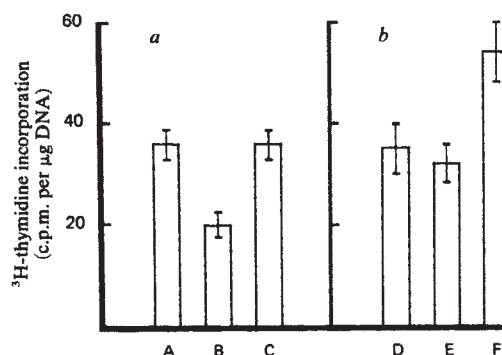


Fig. 2 a, The mitogenic effects of seven daily injections of antisera into 3-week-old intact female rats. Group A received 0.1 ml per day of normal rabbit serum (NRS), group B 0.08 ml anti-N-POC antiserum and group C 0.05 ml of anti-ACTH/LPH antiserum. Six animals were used in each group. Two hours after the last injection, the animals were killed and blood collected for corticosterone estimation. Pairs of adrenals were removed, quartered and incubated in 0.5 ml HeLa medium containing $1 \mu\text{Ci } ^3\text{H}$ -thymidine for 3 h. The glands were then homogenized at 4°C and trichloroacetic acid (TCA) added to a final concentration of 10%. After centrifugation the precipitate was washed three times with ice-cold 10% TCA and then redissolved in 0.3 ml 0.1 M NaOH. Samples were taken for scintillation counting and DNA estimation²⁹ and results expressed as c.p.m. per μg DNA. Group B is significantly different from both the NRS and anti-ACTH/LPH antiserum groups ($P < 0.01$). Corticosterone concentrations (μg per 100 ml) in each group were: A, 15.2 ± 3.1 ; B, 13.4 ± 1.9 ; C, 9.4 ± 1.2 . The anti-ACTH/LPH antiserum was raised in sheep by immunization with ACTH₁₋₂₄ coupled to bovine thyroglobulin by glutaraldehyde. Using radioiodinated ACTH₁₋₃₉ as label, α -MSH, β -MSH, β -LPH and γ -LPH showed full molar cross-reactivity in a radioimmunoassay using this antiserum at a final dilution of 1:12,000. b, The mitogenic effects of N-POC₁₋₇₆ and trypsinized N-POC₁₋₇₆ on the adrenals of 7-week-old female rats. Ten animals were used in each group and dexamethasone blockade was achieved by adding the compound ($10 \mu\text{g ml}^{-1}$) to the drinking water at the start of the experiment. Group E received seven daily injections of $5 \mu\text{g}$ N-POC in $10 \mu\text{l}$ saline. In group F, N-POC₁₋₇₆ (0.5 mg) was treated with DCC-trypsin at a concentration of 1:200 (enzyme/peptide) for 2 h in 1 ml 10 mM NaHCO_3 at 37°C , and was then diluted 10 times with saline and stored at -20°C . Rats received the same amount of peptide as in group E. The saline used for the control group (D) had incubated, heat-treated DCC-trypsin added to it equivalent to that used to digest the injection material in group F. Two hours after the last injection, the animals were killed and the adrenals treated as in a. Thymidine incorporation in group F was significantly higher than in control (D) and intact N-POC₁₋₇₆-treated groups ($P < 0.05$). Corticosterone concentrations (μg per 100 ml) were significantly depressed in all groups: D, 6.8 ± 1.2 ; E, 7.4 ± 0.6 ; F, 8.1 ± 0.8 .

Table 1 Effect of POC peptides on adrenal growth and mitosis *in vivo*

Group	No. rats	Adrenal weight	Mitotic ratio
Control	5	28.08 ± 1.07	0.57 ± 0.21
N-POC ₁₋₂₈	5	$32.5 \pm 0.73^\dagger$	3.39 ± 0.48
N-POC ₁₋₇₆	4	28.85 ± 0.87 NS	0.81 ± 0.32
ACTH ₁₋₂₄	5	$30.86 \pm 1.01^*$	$2.12 \pm 0.35^\ddagger$

Adrenal weight was calculated per 100 g body weight and mitotic ratio was the number of mitoses per mm^2 of adrenal cortex (containing $\sim 4,000$ cells)²⁸. All groups were compared with control values. The difference in mitotic ratio between N-POC₁₋₂₈- and ACTH₁₋₂₄-treated animals was significant ($P < 0.05$). There was no significant variation between the groups with respect to whole body weight. Ten week old female Sprague-Dawley rats were implanted subcutaneously with Alzet osmotic minipumps (model 2001) containing 1 mM acetic acid (control) or 1 mM acetic acid containing $100 \mu\text{g ml}^{-1}$ of the appropriate peptide and delivered to each rat at $\sim 3 \mu\text{g}$ of peptide per day. Rats were distributed randomly between cages and were allowed food and water *ad libitum*. On day 7, each rat was injected with colchicine ($1 \mu\text{g}$ per g body weight) and killed 6 h later. Adrenals were removed, weighed and fixed immediately in Bouin's, prepared for histology in the usual way and stained using Feulgen's method. Three mid-sections of whole cortex of each adrenal (that is, a total of 30 sections per group) were examined microscopically to calculate mitotic ratios. NS, not significant.

* $P < 0.05$; $^\dagger P < 0.02$; $^\ddagger P < 0.01$; $^\S P < 0.001$.

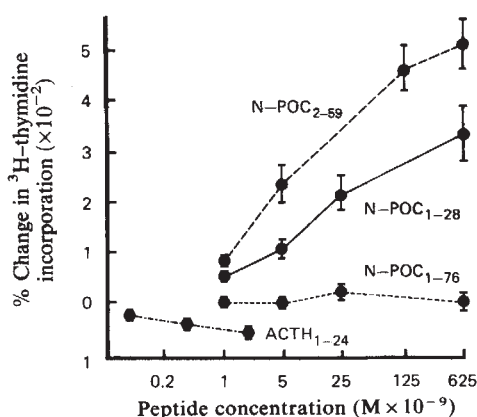


Fig. 3 Dose-response curves of N-POC peptide-mediated stimulation of ^3H -thymidine incorporation into DNA of dispersed rat adrenal cells. Cells were prepared as described previously³⁰. Approximately 10^5 cells were distributed into each vial containing 0.5 ml of HeLa medium containing 0.2% bovine serum albumin (BSA) and 0.2% lima bean trypsin inhibitor (LBI). Cells were preincubated for 3 h before addition of N-POC peptides. After a further 3 h incubation, $1 \mu\text{Ci}$ of ^3H -thymidine was added to each vial and incorporation was allowed to proceed for 6 h. Cells were then centrifuged (100g, 10 min), washed with fresh medium and sonicated in 0.5 ml ice-cold 12% TCA after addition of 100 μg BSA to each vial as carrier. After washing with 10% TCA the precipitate was redissolved in 0.1 M NaOH and treated as for Fig. 2. Control cells gave $2,108 \pm 235$ c.p.m.

(results not shown), both N-POC₁₋₂₈ and N-POC₂₋₅₉, at 10^{-9} – 10^{-7} M, elicited a dose-dependent incorporation of thymidine into DNA, the larger peptide being more potent (Fig. 3). N-POC₁₋₇₆ had little effect at the highest dose assayed and ACTH inhibited thymidine incorporation in a dose-related manner. The synthetic γ -MSHs are very weak steroidogenic agents and N-POC₁₋₂₈ and N-POC₂₋₅₉ are devoid of this activity in the isolated adrenal cell bioassay.

The availability of the smaller N-POC peptides also allowed us to test their adrenal action *in vivo* and to exclude the possibility of artefacts in the thymidine incorporation assay caused by possible changes in the endogenous adrenal thymidine pool size in treated animals. In these experiments we used implanted minipumps to administer the peptides and thus achieved a continuous stimulus over 7 days while completely preserving the endocrine status of the animal. After 1 week's treatment there was no significant difference between the N-POC₁₋₇₆ and control groups with regard to either adrenal weight or mitotic index (Table 1). The N-POC₁₋₂₈-treated group, however, showed a highly significant increase in adrenal weight and mitotic index and seemed more potent than the ACTH₁₋₂₄ group in both respects (Table 1).

Thus although ACTH can be shown to be mitogenic *in vivo* (ref. 1 and Table 1), it initiates a different process of cellular growth *in vivo* and inhibits the hyperplasia seen in the remaining adrenal after unilateral adrenalectomy⁹. It also inhibits cell proliferation *in vitro* (refs 6–8 and Fig. 3). This, coupled with the observations that anti-ACTH antiserum treatment does not lead to adrenal atrophy (ref. 10 and Fig. 2a) or affect the compensatory adrenal hyperplasia in unilateral adrenalectomy¹⁰, argue against ACTH as the physiological mitogen. However, the peptides derived from the N-terminal non- γ -MSH region of proopiomelanocortin used in the present study were active both *in vivo* and *in vitro* and thus may resemble the major adrenal mitogenic hormones. The extensive conservation previously noted in the N-POC region from cDNA sequence data of rat¹⁹, human²⁰ and bovine material is confirmed in a recent comparative immunohistochemical survey, which shows that the anti-N-POC antiserum used in our study immunostains only the corticotropes and melanotropes of a variety of verte-

brate species including fish²¹. This also suggests a biological role for this region.

The need for cleavage of the N-POC region before its mitogenic activity can be expressed suggests an interesting control mechanism. Chemical¹³, biosynthetic and physiological studies²² have indicated that the main molecular form of the corticotropin in blood may be similar to the intact N-POC₁₋₇₆, which is inactive as an adrenal mitogen. The antiserum experiments, however, would suggest either secretion of smaller fragments of N-POC, for example from the pars intermedia (where limited further processing of this region has been shown to take place), or the release of a peptidase which cleaves the peptide after secretion. The most probable cleavage site would be at the double basic sequence at residues 49–50 giving rise to N-POC₁₋₄₈ and N-POC₅₀₋₇₆ (γ -MSH in the ox²³ and rat²⁴ has been found to start with the lysine residue at position 50). Thus the most likely physiological fragment would be N-POC₁₋₄₈. This is corroborated by the observation that the antiserum used in this study effectively removed the naturally circulating biological activity and did not cross-react with smaller biologically active fragments (N-POC₁₋₂₈) in the radioimmunoassay.

In mammals with little or no pars intermedia tissue or its associated peptides, such as humans^{25,26}, post-secretional modification would be the probable mechanism. Thus a change in processing of the N-POC₁₋₇₆ region could provide the necessary mitogenic stimulus to cause, for example, the onset of the adrenarche and development of the androgenic zone of the prepubertal human adrenal gland. As the secretion of peptides derived from N-POC in the corticotrope has been shown to be under the same hypothalamic control as ACTH and β -LPH/endorphin²⁷, their depressed levels during chronic treatment with glucocorticoids could be responsible for the adrenal atrophy observed in such cases. Administration of physiological or even pharmacological doses of peptides derived from this region, whilst continuing glucocorticoid therapy, could provide the necessary mitogenic stimulus and lead to a rapid recovery of the functioning of the hypothalamic pituitary adrenal axis on cessation of that therapy.

Thus we have the unique situation where the N-terminal region of proopiomelanocortin, including ACTH, provides a variety of peptides capable of affecting different aspects of the adrenal gland's activity. N-POC₁₋₂₈, N-POC₂₋₅₉ and ACTH stimulate DNA synthesis, the intact N-POC₁₋₇₆ potentiates the action of ACTH by causing RNA synthesis¹⁵, N-POC₅₁₋₇₆ activates cholesterol ester hydrolase¹⁶ and ACTH switches on the translation of both the stable¹ and N-POC₁₋₇₆-stimulated mRNA pools¹⁰ required for the initiation of corticosteroidogenesis.

The technical assistance of Mrs M. Carino is acknowledged. The work was supported by the Commission de Investigaciones Cientificas de la Provincia de Buenos Aires and by the MRC.

Received 28 October 1981; accepted 2 April 1982.

- Garren, L. D., Gill, G. N., Masui, H. & Walton, G. M. *Recent Prog. Horm. Res.* **27**, 433–474 (1971).
- Dallman, M. F., Engeland, W. C. & Shinsako, J. *Am. J. Physiol.* **231**, 408–414 (1976).
- Gill, G. N., Ill, C. R. & Simonian, M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5569–5573 (1977).
- Payet, N. & Lehoucq, J. G. *J. Steroid Biochem.* **12**, 461–467 (1980).
- Gospodarowicz, D. & Handley, H. H. *Endocrinology* **97**, 102–107 (1975).
- O'Hare, M. J. & Munro Neville, A. *Biochem. Soc. Trans.* **1**, 1088–1091 (1973).
- Masui, H. & Garren, L. D. *Proc. natn. Acad. Sci. U.S.A.* **68**, 3206–3210 (1971).
- Ramachandran, J. & Suyama, A. T. *Proc. natn. Acad. Sci. U.S.A.* **72**, 113–117 (1976).
- Dallman, M. F., Engeland, W. C., Holzwarth, M. A. & Scholz, P. M. *Endocrinology* **107**, 1399–1404 (1980).
- Rao, A. J., Long, J. A. & Ramachandran, J. *Endocrinology* **102**, 371–387 (1978).
- Segal, B. M. & Christy, N. P. *J. clin. Endocr.* **28**, 1465–1472 (1968).
- Nakanishi, S. *et al. Nature* **278**, 423–427 (1979).
- Estivariz, F. E., Hope, J., McLean, C. & Lowry, P. J. *Biochem. J.* **191**, 125–132 (1980).
- Al-Dujaili, E. A. S., Hope, J., Estivariz, F. E., Lowry, P. J. & Edwards, C. R. W. *Nature* **291**, 156–159 (1981).
- Al-Dujaili, E. A. S., Williams, B. C., Edwards, C. R. W., Salacinski, P. R. & Lowry, P. J. *Biochem. J.* **204**, 301–305 (1982).
- Pedersen, R. C., Brownie, A. C. & Ling, N. *Science* **208**, 1044–1045 (1980).
- Hope, J., Ratter, S. J., Estivariz, F. E., McLoughlin, L. & Lowry, P. J. *Clin. Endocr.* **15**, 221–227 (1981).
- McLean, C., Hope, J., Salacinski, P. R., Estivariz, F. E. & Lowry, P. J. *Biosci. Rep.* **1**, 843–849 (1981).
- Drouin, J. & Goodman, H. M. *Nature* **288**, 610–618 (1980).
- Chang, A. C. Y., Cochet, M. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4890–4894 (1980).
- Estivariz, F. E., Iturriza, F. C., Hope, J. & Lowry, P. J. *Gen. comp. Endocr.* **46**, 1–6 (1982).

22. Eipper, B. A. & Mains, R. E. *J. supramolec. Struct.* **8**, 247-262 (1978).
23. Ling, N., Shibasaki, T., Esch, F. & Böhlen, P. 63rd A. Meet. Am. Endocrine Soc. Abstr. no. 383 (The Endocrine Society, 1981).
24. Browne, C. A., Bennett, H. P. J. & Solomon, S. *Biochemistry* **20**, 4538-4546 (1981).
25. Wingstrand, K. G. *The Pituitary Gland* Vol. 3 (eds Harris, G. W. & Donovan, B. T.) 1-27 (Butterworths, London, 1966).
26. Silman, R. E. *et al. Nature* **260**, 716-718 (1976).
27. Estivariz, F. E., Gillies, G. E. & Lowry, P. J. *Pharmac. Ther.* **13**, 61-67 (1981).
28. Van Dorp, A. W. P. & Dean, H. W. *Anat. Rec.* **107**, 265-282 (1950).
29. Giles, K. W. & Myers, A. *Nature* **206**, 93 (1965).
30. Lowry, P. J., McMartin, C. & Peters, J. J. *Endocr.* **59**, 43-55 (1973).

N-methyl-D-aspartate-type receptors mediate striatal ³H-acetylcholine release evoked by excitatory amino acids

Bernard Scatton & John Lehmann*

Synthelabo-LERS, 31 Avenue Paul Vaillant Couturier, 92220 Bagneux, France

Several lines of evidence suggest that striatal cholinergic interneurons receive an excitatory input from the cerebral cortex which utilizes an excitatory amino acid, L-glutamate or L-aspartate, as its neurotransmitter. Cortical ablation reduces striatal high-affinity glutamate uptake^{1,2} and concomitantly decreases acetylcholine turnover³. The destruction of cholinergic, as well as other, neurones of the striatum by the rigid glutamate analogue kainic acid seems to require a functional excitatory amino acid innervation^{4,5}. Furthermore, electron microscopic studies⁶ suggest the existence of a cortical input to the aspiny dendrites of neurones morphologically similar to the presumed striatal cholinergic interneurons^{7,8}. The recent discovery of preferential antagonists of amino acid-induced excitation in the vertebrate central nervous system has led to the classification of excitatory amino acid receptors into three subtypes: the N-methyl-D-aspartate (NMDA), the quisqualate- and the kainate-preferring receptors⁹. We have now attempted to characterize the receptor(s) mediating the excitatory amino acid influence on striatal cholinergic neurones by investigating the effects of specific excitatory amino acid receptor agonists and antagonists on the release of ³H-acetylcholine from slices of the rat corpus striatum. We find that excitatory amino acid analogues evoke a tetrodotoxin-sensitive release of ³H-acetylcholine from rat striatal slices superfused in Mg²⁺-free medium, NMDA and ibotenate being the most potent and kainate and quisqualate the least potent. This effect is antagonized by Mg²⁺ and by (-)-2-amino-7-phosphonoheptanoate (-APHept) and (±)-2-amino-5-phosphonopentanoate (±APPent) but not by glutamic acid diethyl ester (GDEE). These data suggest the involvement of a NMDA-type receptor in the excitatory amino acid influence on striatal acetylcholine release.

The influence of excitatory amino acid receptor agonists on the release of striatal ³H-acetylcholine was first investigated in normal Krebs medium (containing 1.2 mM Mg²⁺). Exposure to 100 µM L-glutamate, L-aspartate and the excitatory amino acid receptor agonists, N-methyl-DL-aspartate, (±)ibotenate, RS-α-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid (AMPA), L-cysteate, kainate and quisqualate of superfused rat striatal slices previously incubated with ³H-choline evoked a moderate release of ³H-acetylcholine (Table 1). AMPA and kainate were the most effective whereas N-methyl-DL-aspartate was almost ineffective. In the absence of calcium, L-glutamate-evoked ³H-acetylcholine release was abolished (data not shown).

As electrophysiological studies on frog and rat spinal neurones have indicated that the different types of excitatory amino acid receptors exhibit a differential sensitivity to Mg²⁺ (low concentrations of Mg²⁺ antagonize NMDA- or L-aspar-

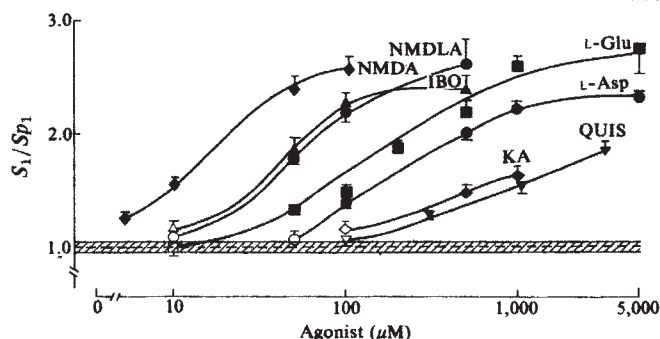


Fig. 1 Effects of excitatory amino acid receptor agonists on the outflow of ³H-acetylcholine from striatal slices in Mg²⁺-free medium: concentration dependence. Methods were as described in Table 1 legend. Each value represents the mean with s.e.m. of eight independent determinations (performed on eight rats). The dotted line and shaded area indicate the basal level with s.e.m. of the release of ³H-acetylcholine in controls ($S_1/Sp_1 = 0.99 \pm 0.05$). Filled symbols: statistically significant ($P < 0.001$) compared with controls (Student's *t*-test). Open symbols: not statistically significant. IBO, (±)ibotenate; NMDA, N-methyl-D-aspartate; NMDLA, N-methyl-DL-aspartate; L-GLU, L-glutamate; L-ASP, L-aspartate; KA, kainate; QUIS, quisqualate.

Table 1 Effects of excitatory amino acid receptor agonists on outflow of ³H-acetylcholine from striatal slices in medium containing 1.2 mM Mg²⁺ and in Mg²⁺-free medium

Agonist (100 µM)	S_1/Sp_1	
	With Mg ²⁺	Without Mg ²⁺
Controls	0.93 ± 0.01	0.94 ± 0.02
N-methyl-DL-aspartate	0.97 ± 0.03	2.12 ± 0.03†‡
(±)Ibotenate	1.05 ± 0.05*	2.07 ± 0.05†‡
AMPA	1.16 ± 0.03†	1.63 ± 0.03†‡
L-glutamate	1.11 ± 0.03†	1.63 ± 0.05†‡
L-aspartate	1.09 ± 0.02†	1.43 ± 0.05†‡
L-cysteate	1.02 ± 0.03*	1.33 ± 0.04†‡
Kainate	1.14 ± 0.03†	1.19 ± 0.02†
Quisqualate	1.03 ± 0.03†	1.06 ± 0.03†

Experiments were performed on male Sprague-Dawley rats (COBS CD Charles River, France, 150 g). Striatal slices (1 × 1 × 0.3 mm prepared by using a McIlwain tissue chopper) were incubated at 37 °C with 0.05 µM [³H]choline (specific activity 80 Ci mmol⁻¹; NEN) in oxygenated Krebs medium (composition in mM: NaCl 118, KCl 4.2, NaHCO₃ 25, NaH₂PO₄ 1.0, CaCl₂ 2.6, glucose 11, MgCl₂ 1.2) for 30 min to allow high-affinity uptake to occur. The slices were subsequently washed, transferred to a Plexiglass chamber (8 mm diameter × 6 mm) and continuously superfused at a constant rate (0.5 ml min⁻¹) with normal or Mg²⁺-free Krebs medium containing 10 µM hemicholinium-3 (a choline uptake inhibitor). The superfusate was collected every 6 min. Excitatory amino acid receptor agonists (100 µM) were added to the superfusion medium for 2 min at 48 min after the beginning of the superfusion. At the end of the experiment, slices were recovered and dissolved in 500 µl 0.5 M sodium hydroxide for radioactivity measurement. Experiments performed in the presence of eserine (30 µM) revealed that the radioactivity released in the medium consisted of 92% authentic ³H-acetylcholine. For each fraction the amount of radioactivity released into the medium as a percentage of total tissue content during that interval (fractional release) was calculated. The elevation of ³H-acetylcholine outflow caused by the agonists is expressed as the ratio of the fractional release obtained in the 6 min interval after the beginning of the 2 min pulse of agonist (S_1) divided by the fractional release obtained in the 6 min period preceding the 2 min pulse of agonist (Sp_1). Each value is the mean with s.e.m. of eight independent determinations (performed on eight animals). The statistical significance of the data was evaluated by the Student's two-tailed *t*-test.

* $P < 0.05$; † $P < 0.01$ compared with controls; ‡ $P < 0.001$ compared with Mg²⁺-containing medium.

* Present address: Department of Neuroscience, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.

tate-induced but not kainate- and quisqualate-induced depolarization)⁹, the responses evoked by excitatory amino acid receptor agonists (100 μ M) were studied in Mg^{2+} -free conditions. When Mg^{2+} was removed from the medium, the ability of *N*-methyl-DL-aspartate to release 3H -acetylcholine increased dramatically (Table 1). Removal of Mg^{2+} also potentiated the response elicited by ($\pm$)ibotenate, AMPA, L-glutamate, L-aspartate and L-cysteate but not that elicited by kainate. This effect was greatest for the agonists which were more specific for NMDA-type receptors (*N*-methyl-DL-aspartate, ibotenate). The relative potencies of excitatory amino acid agonists on the release of 3H -acetylcholine in Mg^{2+} -free medium were compared by plotting concentration-response curves. All the curves were sigmoidal in shape, suggesting the involvement of a single saturable receptor (Fig. 1). The relative potency of the agonists was the following: NMDA > ($\pm$)ibotenate $\approx$ *N*-methyl-DL-aspartate > L-glutamate > L-aspartate > kainate $\approx$ quisqualate. This suggests that the receptor which mediates the release of 3H -acetylcholine evoked by excitatory amino acid agonists shows a preference for NMDA-type agonists.

As suggested by electrophysiological studies, excitatory amino acid receptor subtypes exhibit a differential sensitivity to blockage by various organic antagonists. The most potent and selective NMDA-type receptor antagonists are the phosphonate analogues of carboxylic acids (–)APHept and ($\pm$)APPent, whereas GDEE is considered an antagonist of quisqualate-type receptors^{10,11}. We further characterized the receptor involved in excitatory amino acid-evoked release of 3H -acetylcholine by testing the effects of the amino acid antagonists on the response evoked by *N*-methyl-DL-aspartate and

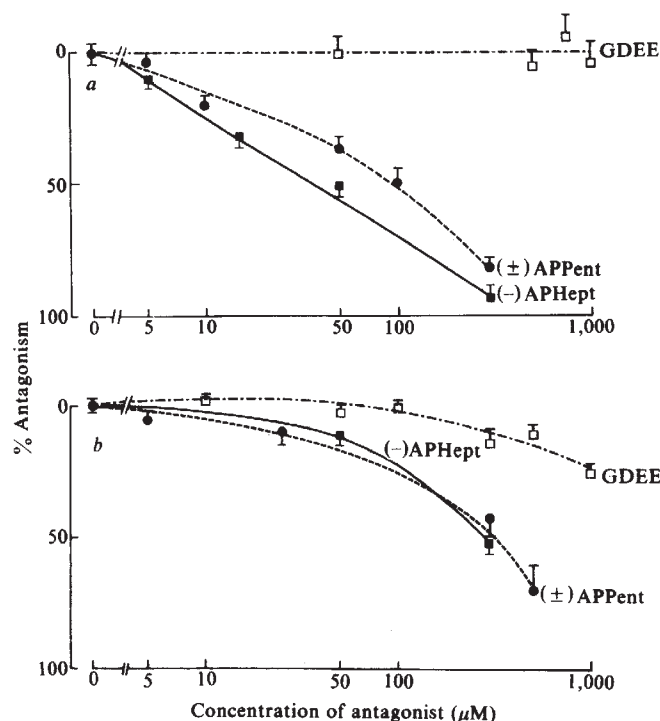


Fig. 2 Effects of excitatory amino acid receptor antagonists on *a*, *N*-methyl-DL-aspartate- and *b*, L-glutamate-evoked release of 3H -acetylcholine from rat striatal slices. Methods were as described in Table 1 legend. The amino acid antagonist in the concentration indicated was added to the superfusion medium 12 min before the 2 min pulse (S_1) of *N*-methyl-DL-aspartate (50 μ M) or L-glutamate (300 μ M). The ratio S_1/S_{P1} was calculated as described in Table 1 legend in both the absence and presence of the antagonist and results were expressed as per cent of this value in respective controls (agonist alone). Results are means with s.e.m. of eight independent determinations (performed on eight animals). Filled symbols: statistically significant (at $P < 0.01$) compared with respective controls (Student's *t*-test). Open symbols: not statistically significant.

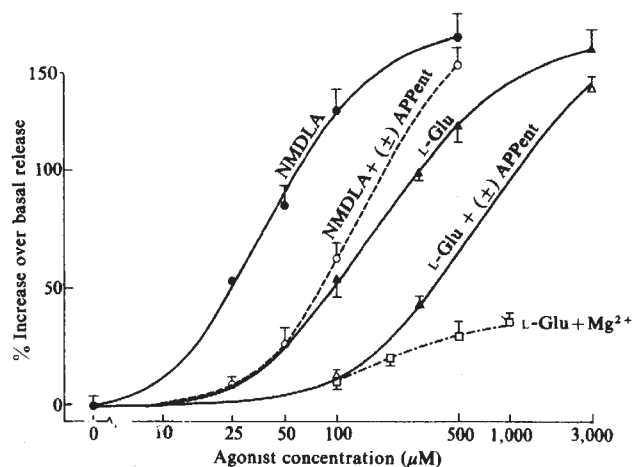


Fig. 3 Competitive antagonism by ($\pm$)APPent and non-competitive antagonism by Mg^{2+} of the *N*-methyl-DL-aspartate- and L-glutamate-evoked release of striatal 3H -acetylcholine. For methods and expression of results see Table 1 legend. Mg^{2+} (1.2 mM) or ($\pm$)APPent (50 and 300 μ M when tested against NMDLA and L-glutamate respectively) were introduced in the superfusion medium 12 min before the 2 min pulse of the excitatory amino acid agonist. Each point corresponds to the mean with s.e.m. of eight independent determinations (performed on eight animals).

L-glutamate in Mg^{2+} -free conditions. (–)APHept and ($\pm$)APPent antagonized in a concentration-dependent manner the release of 3H -acetylcholine evoked by *N*-methyl-DL-aspartate (50 μ M) with IC_{50} s of 40 and 90 μ M respectively (Fig. 2). Both drugs also antagonized the response to L-glutamate (300 μ M) with IC_{50} s of 280 and 320 μ M respectively. In contrast, the quisqualate receptor antagonist GDEE failed to affect the response evoked by *N*-methyl-DL-aspartate and only weakly antagonized (as a concentration of 1 mM) the L-glutamate response. The inhibition by ($\pm$)APPent of the *N*-methyl-DL-aspartate- and L-glutamate-evoked release of 3H -acetylcholine is clearly competitive (Fig. 3), in contrast to the Mg^{2+} inhibition which seems to be non-competitive. In the concentration range tested, neither (–)APHept nor ($\pm$)APPent alone affected the spontaneous outflow of 3H -acetylcholine. Moreover, at a concentration effective in blocking *N*-methyl-DL-aspartate or L-glutamate responses, ($\pm$)APPent only marginally inhibited the release of 3H -acetylcholine evoked by other depolarizing agents— K^+ (12 mM) and veratridine (1 μ M) S_1/S_{P1} : controls, 0.96 ± 0.02 ; K^+ , 2.74 ± 0.06 ; K^+ + ($\pm$)APPent, 2.30 ± 0.01 , $P < 0.01$, Student's two-tailed *t*-test; veratridine, 1.70 ± 0.09 ; veratridine + ($\pm$)APPent, 1.61 ± 0.04 , $n = 8$), suggesting that the antagonism occurs at the level of specific amino acid receptors.

The present data are consistent with the involvement of an NMDA-type receptor in the excitatory amino acid influence on striatal 3H -acetylcholine release. This conclusion is supported by (1) the high potency of the NMDA-type receptor agonists NMDA and ($\pm$)ibotenate to evoke the release of the labelled transmitter in Mg^{2+} -free conditions; (2) the antagonism by low concentrations of Mg^{2+} of NMDA-type agonist-evoked release of striatal 3H -acetylcholine; (3) the selective blockage by antagonists of NMDA-type receptors of the *N*-methyl-DL-aspartate- and L-glutamate-evoked release of 3H -acetylcholine in Mg^{2+} -free medium. Note that NMDA has also been found to be the most potent excitatory amino acid in stimulating sodium efflux from striatal slices¹². The very weak potency of quisqualate, together with the failure of the quisqualate-type antagonist GDEE to block *N*-methyl-DL-aspartate- or L-glutamate-evoked 3H -acetylcholine release in both Mg^{2+} -free conditions and Mg^{2+} -containing medium (not shown), suggest that quisqualate-type receptors do not contribute significantly to the excitatory amino acid-evoked responses of striatal cholinergic cells.

The localization of the NMDA receptors mediating ^3H -acetylcholine release evoked by excitatory amino acid agonists is not yet known. However, the observation that addition of tetrodotoxin ($0.5\ \mu\text{M}$) abolished the *N*-methyl-DL-aspartate ($50\ \mu\text{M}$) and L-glutamate ($300\ \mu\text{M}$)-evoked release of ^3H -acetylcholine in Mg^{2+} -free medium (S_1/S_{P1} : controls, 0.96 ± 0.02 ; *N*-methyl-DL-aspartate, 2.88 ± 0.06 ; *N*-methyl-DL-aspartate + tetrodotoxin, $1.10 \pm 0.07^*$; L-glutamate, 2.75 ± 0.02 ; L-glutamate + tetrodotoxin, $1.09 \pm 0.03^*$; $*P < 0.01$, Student's *t*-test, compared with agonist alone) indicates that action potentials mediate the release of ^3H -acetylcholine evoked by these agonists. These data are consistent with the hypothesis that excitatory amino acid receptor agonists activate receptors located on cholinergic dendrites, causing depolarization and action potential propagation along the axons of the cholinergic interneurons to the terminals, where ^3H -acetylcho-

line is released. The involvement of excitatory amino acid receptors on cells other than cholinergic neurones which affect the cholinergic neurones via an unknown transmitter cannot, however, be excluded rigorously.

In conclusion, an NMDA-type receptor has been found to mediate the excitatory influence of amino acids on striatal cholinergic neurones. These receptors, which may be located on dendrites of cholinergic cells, may play a part in transducing information received from the cerebral cortical afferents, which are thought to use an excitatory amino acid as neurotransmitter.

We thank Dr J. F. Collins, P. Krogsgaard-Larsen and J. C. Watkins for gifts of compounds, and D. Fage for technical assistance. (–)APHept was obtained from Dr Collins, Department of Chemistry, City of London Polytechnic, and (±)APPent from Dr Watkins, Department of Pharmacology and Physiology, The Medical School, Bristol.

Received 29 January; accepted 14 April 1982.

- McGeer, P. L., McGeer, E. G., Scherer, U. & Singh, K. *Brain Res.* **128**, 369–373 (1977).
- Divac, I., Fonnum, F. & Storm-Mathisen, J. *Nature* **266**, 377–378 (1977).
- Wood, P. L., Moroni, F., Cheney, D. L. & Costa, E. *Neurosci. Lett.* **12**, 349–354 (1979).
- McGeer, E. G., McGeer, P. L. & Singh, K. *Brain Res.* **139**, 381–383 (1978).
- Biziere, K. & Coyle, J. T. *Neurosci. Lett.* **8**, 303–310 (1978).
- Hassler, R., Chung, J. W., Rinne, U. & Wagner, A. *Expl Brain Res.* **31**, 67–80 (1978).
- Lehmann, J. & Fibiger, H. C. *Life Sci.* **25**, 1939–1947 (1979).

- Kimura, H., McGeer, P. L., Peng, F. & McGeer, E. G. *Science* **208**, 1057–1059 (1980).
- Watkins, J. C. in *Glutamate: Transmitter in the Central Nervous System* (eds Roberts, P. J., Storm-Mathisen, J. & Johnston, G. A. R.) 1–24 (Wiley, Chichester, 1981).
- Watkins, J. C., Davies, J., Evans, R. H., Francis, A. A. & Jones, A. W. in *Glutamate as a Neurotransmitter* (eds Di Chiara, G. & Gessa, G. L.) 263–273 (Raven, New York, 1981).
- Perkins, M. N., Stone, T. W., Collins, J. F. & Curry, K. *Neurosci. Lett.* **23**, 333–336 (1981).
- Luini, A., Goldberg, O. & Teichberg, V. I. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3250–3254 (1981).

Modulation of lateral mobility of band 3 in the red cell membrane by oxidative cross-linking of spectrin

Deborah K. Smith & Jiri Palek

The Department of Research, St Elizabeth's Hospital, Tufts University School of Medicine, Boston, Massachusetts 02135, USA

Several recent studies have shown that the lateral diffusion of erythrocyte transmembrane proteins is modulated by the sub-membrane network of spectrin, actin and polypeptide 4.1 (refs 1–3). The rate of lateral diffusion of the major transmembrane glycoprotein (the anion transport channel, band 3) is increased approximately 50-fold in membranes from spherocytic erythrocytes of mice which are deficient in the membrane skeletal protein spectrin¹. Lateral mobility is increased when spectrin is removed from the membrane by extraction at low ionic strength² or when it is displaced by a fragment of ankyrin containing the membrane attachment site for spectrin³. Such experiments show that the presence of spectrin on the membrane restricts the mobility of membrane proteins, but they do not define the mechanism by which spectrin could modulate membrane protein mobility *in situ*. The oxidant diamide has been shown to cross-link spectrin via intermolecular disulphide coupling into a high molecular weight complex⁴ and this is associated with inhibition of discocyte–echinocyte shape change⁵ and decreased erythrocyte deformability⁶. Here we report that the lateral mobility of band 3 protein is decreased in erythrocytes in which spectrin is cross-linked via diamide. Furthermore, this inhibition of mobility is reversed when spectrin oxidative cross-links are reduced by addition of dithiothreitol (DTT).

Membranes isolated from red cells treated with 2 mM diamide or diamide followed by DTT were subjected to SDS-polyacrylamide gel electrophoresis in non-reducing conditions (Fig. 1). Membranes of cells treated with diamide showed the presence of a high molecular weight polymer ($>10^6$ molecular weight) which was reducible by addition of DTT to the gel. The polymer was absent in cells treated with DTT, indicating

that it was also reversible in intact cells. Diamide-treated samples were electrophoresed in the second dimension in reducing conditions to identify components of the high molecular weight polymer. As shown in Figure 2, the polymer consisted primarily of spectrin ($>90\%$) with small amounts of other membrane proteins. There was some spectrin dimer and some band 3 dimer, which might be expected because these proteins probably exist as tetramers in the membrane^{7,8}. Cross-linked spectrin dimers were routinely detected in heavily loaded non-reducing gels, such as in the first-dimension gel of Fig. 2. There was no evidence, however, of cross-linking of band 3 to other membrane components such as ankyrin or spectrin, which might affect band 3 mobility directly. Therefore, in these conditions and in intact cells, diamide cross-linked spectrin fairly specifically. More extensive treatment with diamide ($>2\ \text{mM}$)

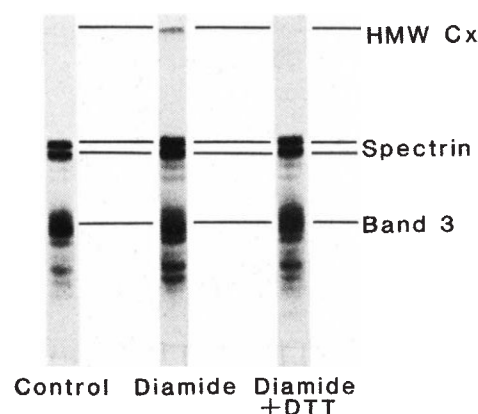


Fig. 1 Non-reducing SDS-polyacrylamide gels of membranes from diamide-treated erythrocytes. Blood was collected from healthy volunteers. Cells were washed in saline and buffy coats removed. Erythrocytes were incubated with 2 mM diamide in 5 mM Na_2HPO_4 , 145 mM NaCl, pH 7.4 for 30 min at 37°C . Cells were washed with the same buffer, and an aliquot of diamide-treated cells was incubated for 30 min at 37°C with 15 mM DTT. Cells were washed, membrane ghosts prepared and samples were electrophoresed on 2.5% acrylamide–1% agarose gels as described previously¹³. All gels were stained with Coomassie brilliant blue. HMW Cx., high molecular weight complex.

caused significant cross-linking of other membrane proteins as well as spectrin (not shown).

To determine band 3 lateral diffusion rates, erythrocytes were labelled with the fluorescent molecule 5-(3,5-dichlorotriazinyl) aminofluorescein (DCTAF) in conditions where band 3 protein was specifically labelled¹. Labelled cells were then fused with unlabelled cells by the addition of polyethylene glycol, and fused red cell doublets were observed in the fluorescence microscope for spread of fluorescence from the labelled to the unlabelled half. Label equilibrated significantly more slowly in erythrocytes in which spectrin was oxidatively cross-linked with diamide (Fig. 3). The time necessary for equilibration of label on 90% of all doublets (Fig. 3) was used to calculate a minimum diffusion constant ($D_{\min}$) for band 3 according to the procedures of Fowler and Branton⁹ and Huang¹⁰. $D_{\min}$ for diffusion of band 3 in normal red cell membranes was calculated to be $3 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$, which agrees with previously published values obtained by this and other methods^{1-3,9,11}. $D_{\min}$ for diamide-treated cells was $\sim 1.8 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$, reflecting a 40% reduction in the rate of band 3 diffusion ($P < 0.005$). In diamide-treated cells which were further incubated with DTT to reduce oxidative cross-links, $D_{\min}$ was the same as for untreated cells. In experiments where DTT incompletely reduced oxidative cross-links, the extent of reversal of diamide effects on mobility were proportional to the degree of DTT reduction of the high molecular weight membrane polymer formed by diamide (not shown). These results indicate that stabilization of spectrin in a high molecular weight complex slows band 3 diffusion without cross-linking band 3 to the membrane skeleton.

Membranes isolated from red cells after fusion and equilibration of label showed no proteolysis when analysed on SDS-acrylamide gels. Furthermore, addition of the protease inhibitors phenylmethylsulphonyl fluoride (1 mM), *N*- α -tosyl-L-lysine chloromethyl ketone (1 mM) and pepstatin A ($1 \mu\text{g ml}^{-1}$) had no effect on mobility of band 3 in these experiments.

Previous studies showed that the membrane skeleton serves to restrict lateral movement of transmembrane proteins, and two mechanisms have been proposed. First, studies by Golan and Veatch² and by Fowler and Bennett³ on the dissociation of spectrin from the membrane suggested that the diffusion rate

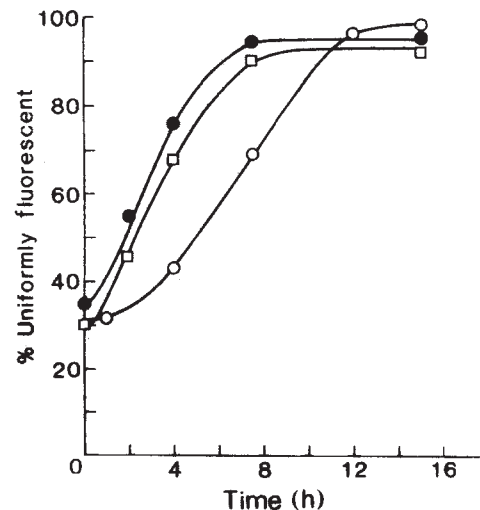


Fig. 3 Effect of diamide-induced oxidation of spectrin on lateral mobility of band 3 protein at 30°C. Erythrocytes at 20% haematocrit in 10 mM NaHCO_3 , 145 mM NaCl, pH 9.5, were incubated on ice for 1 h with 2 mg ml^{-1} DCTAF. Cells were washed twice in 10 mM NaHCO_3 , 95 mM NaCl, 50 mM glycine, pH 9.5 to remove unreacted DCTAF, then washed twice in phosphate-buffered saline (10 mM Na_2HPO_4 , 145 mM NaCl, pH 7.4). Unlabelled cells were incubated with the same buffers in the absence of DCTAF. The labelling procedure caused no haemolysis. Labelled and unlabelled cells were fused with polyethylene glycol as described earlier¹¹. Fused doublets were observed at 30°C for diffusion and equilibration of label over the entire doublet. Data are plotted as per cent of labelled doublets which are uniformly fluorescent versus time⁹. ○, Cells treated with 2 mM diamide at 37°C for 30 min before labelling and fusion. ●, Cells treated with diamide, washed, then incubated with 15 mM DTT at 37°C for 30 min before labelling and fusion. □, Cells incubated at 37°C with no diamide or DTT before labelling or fusion.

of band 3 can be controlled by reversible interaction of band 3 with spectrin via band 2.1. Second, Cherry *et al.* proposed that spectrin forms a meshwork underneath the membrane, with interstices through which band 3 can move¹². Thus, lateral diffusion of transmembrane proteins could be modulated by the state of the spectrin meshwork itself, independently of spectrin-band 3 interactions. The latter hypothesis was supported by observations of Schindler *et al.* that agents which disrupt cytoskeletal structure also free band 3 to diffuse laterally in the membrane¹¹. Our results are also consistent with this model of the membrane skeleton as a dynamic structure that modulates lateral movement of membrane proteins. When this dynamic state is stabilized, for example by cross-linking with spectrin, mobility of transmembrane proteins is reduced.

This work was supported by NIH grant HL15157.

Received 29 December 1981; accepted 31 March 1982.

1. Sheetz, M. P., Schindler, M. & Koppel, D. E. *Nature* **285**, 510–512 (1980).
2. Golan, D. E. & Veatch, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2537–2541 (1980).
3. Fowler, V. & Bennett, V. *J. supramolec. Struct.* **8**, 215–221 (1978).
4. Haest, C. W. M., Kamp, D., Plasa, G. & Deuticke, B. *Biochim. biophys. Acta* **469**, 226–230 (1977).
5. Haest, C. W. M., Fischer, T. M., Plasa, G. & Deuticke, B. *Blood Cells* **6**, 539–553 (1980).
6. Fischer, T. M., Haest, C. W. M., Stohr, M., Kamp, D. & Deuticke, B. *Biochim. biophys. Acta* **510**, 270–282 (1978).
7. Palek, J. & Liu, S.-C. *Red Cell Proc. 5th int. Conf. on Red Cell Metabolism & Function* (ed. Brewer, G.) (Liss, New York, in the press).
8. Weinstein, R. S., Khodadad, J. K. & Steck, T. L. *J. Cell Biol.* **87**, 209a (1980).
9. Fowler, V. & Branton, D. *Nature* **268**, 23–26 (1977).
10. Huang, H. W. *J. theor. Biol.* **40**, 11–17 (1973).
11. Schindler, M., Koppel, D. E. & Sheetz, M. P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1457–1461 (1980).
12. Cherry, R. J., Burkli, A., Busslinger, M., Schneider, G. & Parish, G. R. *Nature* **263**, 389–393 (1976).
13. Liu, S. C., Fairbanks, G. & Palek, J. *Biochemistry* **16**, 4066–4074 (1977).
14. Laemmli, U. K. *Nature* **227**, 680 (1970).

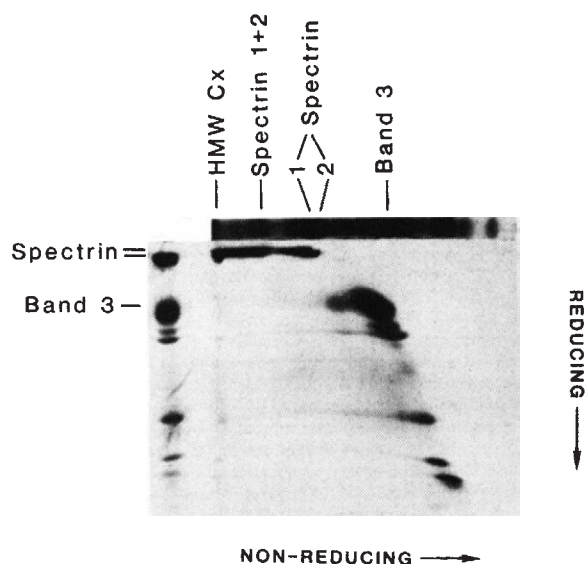


Fig. 2 Two-dimensional SDS-polyacrylamide gel electrophoresis of membranes from diamide-treated erythrocytes. Erythrocytes were diamide-treated, membrane ghosts prepared and membranes were electrophoresed in the first non-reducing dimension as in Fig. 1. This gel was then overlaid on a 7% acrylamide slab gel and electrophoresed in reducing conditions¹⁴.

Higher plant tubulin identified by self-assembly into microtubules *in vitro*

Louis C. Morejohn & Donald E. Fosket

Department of Developmental and Cell Biology,
University of California, Irvine, California 92717, USA

Microtubules are filamentous, subcellular structures present in virtually all eukaryotes. In higher plants, microtubules form the mitotic spindle¹, determine the plane of cell division^{2,3} and orient cellulose microfibril deposition in growing cells^{4,5}, and thus are significant determinants of morphogenesis^{6,7}. The main structural component of microtubules is tubulin, a conserved, heterodimeric protein having a molecular weight of 110,000 and composed of α - and β -subunits ($\approx 55,000$ each)⁸. The resistance of plant microtubules to several antimitotic drugs has suggested that plant tubulins differ from animal tubulins (see refs 5, 9, 10 for review), but further characterization has awaited the development of a method to isolate plant tubulin. As tubulin has no characteristic enzymatic activity, the positive identification of this protein depends on its ability to self-assemble *in vitro* to form structures which have the morphological features of microtubules. Recent studies¹¹⁻¹⁷ have used indirect methods to isolate tubulin-like proteins from higher plant cell extracts, but none has conclusively identified tubulin by demonstrating its self-assembly into microtubules. We report here the isolation of tubulin from cultured cells of a higher plant and its unequivocal identification by self-assembly into microtubules *in vitro*. Furthermore, we report peptide mapping data which indicate that although the β -subunits of mammalian and higher plant tubulins have been conserved, the α -subunits have diverged during evolution.

Previous work from our laboratory demonstrated that tubulin self-assembly in extracts of Paul's Scarlet rose (PSR) cells is hampered by low tubulin concentrations, by low molecular weight inhibitors and by endogenous peptidases¹⁷. We have circumvented these problems by designing a tubulin isolation method that is particularly suited to the physiological characteristics of PSR cells grown in suspension. Early exponential phase (day 3-5) cells were used because during this period the concentrations of cellular proteins are highest¹⁸ and those of potentially harmful polyphenolic compounds lowest¹⁹. Extracts were prepared in a buffer designed to minimize proteolysis and to enzymatically cleave nucleic acids which, as polyanions, are known to inhibit *in vitro* assembly of tubulin²⁰. In addition, anion-exchange chromatography was used to select and concentrate tubulin from extracts by virtue of its conserved polyanionic charge²¹.

PSR cells, originally isolated from rose stem tissues (*Rosa* sp. cv. Paul's scarlet) by Nickell and Tulecke²², were grown as stock suspension cultures in a defined medium²³, modified as previously described¹⁷. For the propagation of cells to be used for protein isolation, aliquots of 14-day-old stationary phase cells were diluted 6.6-fold with fresh medium and cultured for 3-5 days.

Extracts of cultured cells were prepared, and proteins were fractionated on DEAE-Sephadex A50 as modified from Weisenberg *et al.*²⁴ (see Fig. 1 legend) and analysed by SDS-polyacrylamide gel electrophoresis²⁵ (SDS-PAGE). Step gradient elution of the column separated the crude supernatant into three protein-containing fractions designated A, B and C (Fig. 1). SDS-PAGE of these fractions demonstrated that most of the cellular protein eluted in fractions A and B, while fraction C was enriched for two polypeptides with mobilities similar to those of the α - and β -subunits of bovine brain tubulin. Densitometer tracings of stained gels revealed that these bands comprised $\sim 63\%$ of the fraction C polypeptides that entered the gel. Thus DEAE chromatography represents a significant step in the purification of tubulin-like polypeptides from a

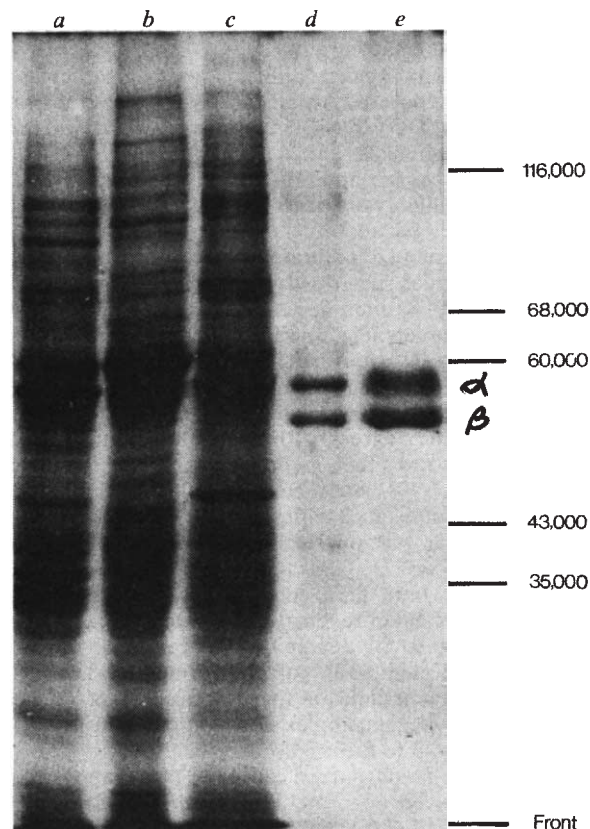


Fig. 1 Analysis of fractionated PSR proteins by SDS-PAGE²⁵. Early exponential phase PSR suspension cultures were placed in a dark cold room at 4 °C with constant shaking for 2 h before collection on a Miracloth filter, washed with cold PM buffer (50 mM PIPES-KOH pH 6.9, 1 mM EGTA, 0.5 mM MgCl₂, 1 mM dithiothreitol). The cells were then weighed and homogenized in a chilled Sorval Omni-mixer in PM buffer containing 2 mM GTP (5 g cells per ml buffer) for 3 min. The homogenate was filtered through three layers of Miracloth and the filtrate was made 25 μ M in leupeptin hemisulphate (Sigma) and 100 μ g ml⁻¹ in both DNase I (Sigma) and RNase (Calbiochem). The filtered homogenate was centrifuged at 48,200 g for 30 min at 4 °C. The crude supernatant, which contained 2.4-3.3 mg ml⁻¹ of protein as assayed by the method of BioRad²⁶, was poured into a flask containing DEAE-Sephadex A50 (ref. 24) (pre-equilibrated with PM buffer) at a ratio of 0.5 ml packed volume of resin per ml of crude supernatant. The mixture was swirled in the dark at 4 °C for 1 h, briefly de-aerated and poured into a 2.6 $\times$ 40 cm column. The column was packed with crude supernatant at a flow rate of 2-3 ml min⁻¹, and unbound proteins (fraction A) were eluted directly. Weakly bound proteins (fraction B) were eluted with 3-5 bed volumes of PM buffer containing 0.4 M KCl and 0.5 mM GTP. Tightly bound proteins (fraction C) were eluted from the column with 3 bed volumes of PM buffer containing 0.8 M KCl and 0.5 mM GTP. Each fraction was collected by precipitation by 50% saturation with (NH₄)₂SO₄ for 30 min. Immediately prior to electrophoresis, (NH₄)₂SO₄ precipitates were dissolved in PM buffer, desalted on Sephadex G25 columns and acetone-precipitated, and the precipitates were boiled in SDS-sample buffer²⁵ for 2 min. Crude supernatant samples were boiled directly in equal volumes of SDS-sample buffer. SDS-derivatives were loaded on to 9.25% slab gels prepared according to Studier²⁵, except that the stacking gel was 5.1% acrylamide, the resolving gel contained 0.15 M Tris-HCl pH 8.8, and the SDS (Sigma) was ethanol-recrystallized. Gels were stained overnight with 0.025% Coomassie brilliant blue R-250, destained, dried and photographed with a yellow filter. a, PSR crude supernatant polypeptides; b, fraction A polypeptides; c, fraction B polypeptides; d, fraction C polypeptides containing putative tubulin α - and β -subunits; e, three times assembled bovine brain tubulin α - and β -subunits prepared according to ref. 33. Molecular weight standards were *Escherichia coli* β -galactosidase (116,000), bovine serum albumin (68,000), bovine liver catalase (60,000), ovalbumin (43,000) and bovine heart lactate dehydrogenase (35,000).

heterologous population of proteins in PSR cell extracts. The yield of presumptive tubulin obtained by this procedure was 1-2 μ g mg⁻¹ of crude supernatant proteins as determined by BioRad assays²⁶.

The proteins in fraction C were tested for their ability to form microtubules *in vitro*. Assembly reactions were performed following the addition of glycerol²⁷ or taxol²⁸, an experimental

antitumour drug (Fig. 2 legend). Electron microscopy of the materials assembled in the presence of taxol revealed numerous short microtubules, while materials formed in the presence of glycerol contained microtubules which were less abundant and usually longer (Fig. 2). Protein determinations²⁶ on the microtubule pellets showed that ~7% of the total protein present in fraction C formed pelletable aggregates in the glycerol-treated sample, whereas ~44% of the total protein was pelleted from the taxol-treated sample.

Analysis of the *in vitro*-assembled materials by SDS-PAGE²⁵ showed that tubulin was the principal protein present in the pelleted aggregates assembled in the presence of both glycerol and taxol (Fig. 3). Furthermore, the rose tubulin polypeptides had mobilities similar to those of bovine brain tubulin polypeptides. In this particular gel system, the β -subunits co-migrated and the rose α -subunit had a slightly faster mobility than that of the brain α -subunit (compare subunits in lanes *d*, *e* of Fig. 3). Similarly, the α -subunits of flagellar tubulins from the green alga *Chlamydomonas* and the ferns *Marselia* and *Pteridium* have been shown to exhibit electrophoretic mobilities different from that of mammalian α -tubulin^{29,30}.

Densitometer tracings of electrophoretically separated microtubule pellets showed that tubulin represented ~36% of

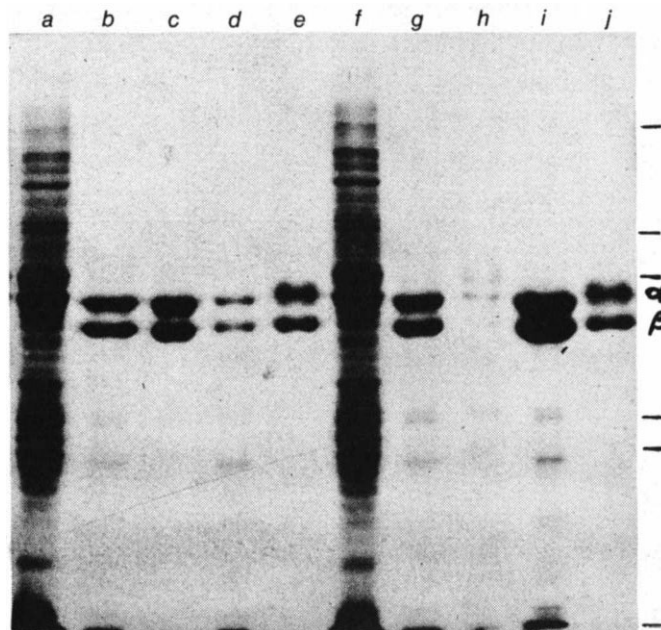


Fig. 3 Analysis of microtubule assembly mixtures by SDS-PAGE²⁵. The supernatant and pellet samples described in Fig. 2 legend were boiled in equal volumes of SDS-sample buffer²⁵ and loaded on to the gel in equal volumes. In this way, we could compare directly the polypeptide composition and abundance of tubulin in each fraction from the glycerol (lanes *c* and *d*) and taxol (lanes *h* and *i*) treatments. *a*, *f*, PSR crude supernatant polypeptides; *b*, *g*, fraction C polypeptides; *c*, glycerol-treated supernatant polypeptides; *d*, glycerol-treated microtubule pellet polypeptides; *e*, *j*, three times assembled bovine brain tubulin³³; *h*, taxol-treated supernatant polypeptides; *i*, taxol-treated microtubule pellet polypeptides. Molecular weight markers as in Fig. 1.

the total protein in the glycerol-assembled material, and ~74% of the taxol-microtubule pellet. Apparently a single round of assembly did not enrich for tubulin in the glycerol-treated samples but did so substantially in the presence of taxol. Horwitz and co-workers have shown that taxol enhances both the assembly rate and yield of brain microtubules formed *in vitro*^{28,31}. They have further demonstrated that taxol cannot be readily removed from microtubules, and that it stabilizes the polymers against depolymerization by cold treatment, CaCl_2 , podophyllotoxin or vinblastine³². For these reasons taxol-induced microtubules cannot be further purified by successive rounds of polymerization. Also, the low yield of tubulin from PSR cells and the low efficiency of glycerol-induced microtubule assembly have prohibited the use of cyclic assembly procedures³³. We are now investigating methods to improve the yield of tubulin from these cells.

Previous studies on ion-exchange-purified mammalian brain tubulin have shown the *in vitro* assembly reaction to be enhanced by both glycerol²⁷ and taxol³¹. The present study demonstrates that plant tubulin isolated by DEAE chromatography assembles into microtubules after addition of glycerol or taxol. Both the number of microtubules and the yield of tubulin increased when taxol was present. Similar results have been obtained in studies on taxol-induced mammalian tubulin assembly^{28,31}, and so it seems that the tubulins from organisms as diverse as mammals and higher plants are subject to similar constraints in terms of their abilities to assemble into microtubules *in vitro*.

The microtubules in algae, fungi, ferns and higher plants show resistance to several antimicrotubule agents commonly used to disrupt animal microtubules, suggesting significant differences between plant and animal tubulins^{5,10}. Indeed, recent studies have revealed several differences between the tubulins from some of these organisms and from mammals, including variations in electrophoretic mobility^{29,34,35}, sensitivity to drugs during *in vitro* assembly³⁵⁻³⁷ and peptide mapping patterns^{30,34}.

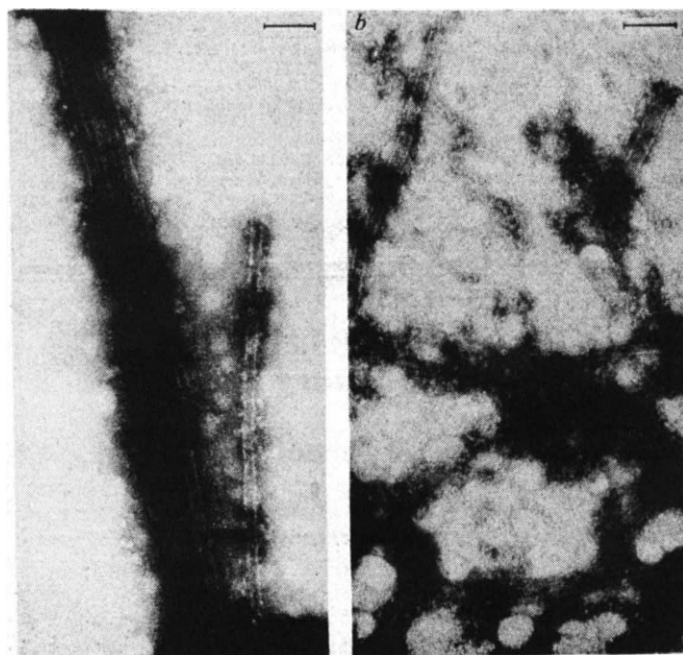


Fig. 2 Electron micrographs of negatively stained rose microtubules assembled *in vitro*. For *in vitro* assembly studies, the fraction C $(\text{NH}_4)_2\text{SO}_4$ precipitates (described in Fig. 1 legend) from ~200 g of cells were dissolved in 1–1.5 ml of PM buffer containing 2 mM GTP. The solution was desalted on a Sephadex G25 column (5 ml bed volume) which had been pre-equilibrated with the same buffer. The excluded volume was made 50 μM in leupeptin hemisulphate and concentrated by dialysis against 40 ml of the above buffer containing 8 M glycerol and 50 μM leupeptin hemisulphate at 4 °C for 2 h. This reduced the volume of the sample ~10-fold to give a 100–200 μl sample containing 2.6–3.6 mg ml^{-1} of protein as determined by the BioRad²⁶ procedure. This sample was tested for its ability to form microtubules after addition of glycerol or taxol. The sample was divided equally into two samples, one of which was made 4 M in glycerol²⁷, while the other was made 10 μM in taxol²⁸. The samples were incubated at 37 °C for 1 h to permit microtubule assembly and centrifuged at 130,000 g in a Beckman airfuge for 1 h at 23 °C. Each supernatant was removed and divided into two aliquots, one to be used for protein determination and the other for SDS-gels. Each microtubule pellet was resuspended in 10 μl of PM buffer containing 4 M glycerol. One μl of the suspension was removed for analysis by electron microscopy, while the remaining suspension was divided into two aliquots, one to be used for protein determinations and the other for SDS-gels. Samples for electron microscopy were placed on carbon- and formvar-coated copper grids, displaced with water, stained with 1% uranyl acetate for 30 s and air-dried. Samples were photographed on a Zeiss 9S electron microscope. *a*, Microtubules assembled in glycerol-treated sample; *b*, microtubules assembled in taxol-treated sample. Bar, 0.1 μm .

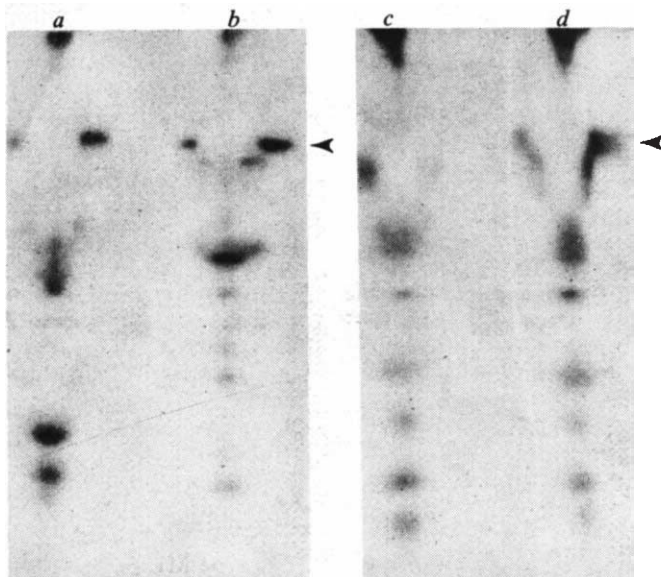


Fig. 4 Peptide mapping analysis of the α - and β -subunits of bovine brain tubulin and rose tubulin. Bovine brain tubulin purified by three rounds of assembly³³ and rose tubulin purified by taxol-induced assembly were analysed by limited proteolysis in gels according to the method of Cleveland *et al.*³⁸, with minor modifications. The subunits of these tubulins were first separated on the gel system described in Fig. 1 legend. Gels were stained with 0.2% Coomassie brilliant blue R-250 in 50% methanol for 30 min and destained in 5% methanol for 15 min. Bands corresponding to the α - and β -subunits were cut out, equilibrated in a solution of 0.125 M Tris-HCl pH 6.8, 0.1% SDS and 1 mM EDTA for 1 h, and loaded on to 15% Laemmli gels³⁹ together with 0.5 μ g of *Staphylococcus aureus* V8 protease. Electrophoresis was performed according to Cleveland *et al.*³⁸. Gels were stained and processed as described for Fig. 1. Lanes contain polypeptide fragments from a, bovine brain α -subunit; b, rose α -subunit; c, bovine brain β -subunit; d, rose β -subunit. Arrows indicate positions of uncleaved subunits.

We have started structural analysis of cytoplasmic tubulin from a higher plant by preparing peptide maps³⁸ of the α - and β -subunits of rose tubulin and bovine brain tubulin (Fig. 4). Cleavage of the β -subunits of both tubulins by *Staphylococcus aureus* V8 protease produced very similar patterns of polypeptide fragments. In contrast, the patterns of peptide fragments generated from the cleavage of the α -subunit of both tubulins not only differed markedly from each other, but also from the respective β -subunit patterns. These data suggest that the β -subunits of plants and animals have been more conserved during evolution than the α -subunits.

Recently published³⁰ peptide maps of the tubulin subunits from cow, squid, sea urchin, echinurid worm, a fern and an alga show the β -subunit maps of all these organisms to be quite similar, the α -subunit maps less similar and the plant α -subunit maps markedly different from those of the animals. Moreover, the fern and algal α -subunit maps resemble the α -subunit map of slime mold tubulin reported by Clayton *et al.*³⁴. Our peptide mapping data on higher plant tubulin are consistent with the idea³⁴ that the α -subunits of tubulins from organisms that are generally resistant to anti-microtubule agents have diverged with respect to the α -subunits from organisms sensitive to those agents. Experiments are in progress to determine whether the resistance of higher plant cells to anti-microtubule drugs is, in fact, due to a low affinity of plant tubulin for these drugs, or to other modes of resistance¹⁰.

We are raising antibodies to plant tubulin subunits for use in further studies on these polypeptides. Efforts are also underway to construct specific cDNA probes for use in studies on the regulation of synthesis of tubulin mRNAs in higher plants.

We thank Dr Susan Horwitz for her generous gift of taxol and Drs Marc Kirchner and Timothy Bradley for helpful suggestions during this investigation. This research was supported by grants to D.E.F. from the NSF (7722398) and the Monsanto Chemical Co., St Louis, Missouri.

Received 5 March; accepted 2 April 1982.

- Nicklas, R. B. in *Molecules and Cell Movement* (Raven, New York, 1975).
- Pickett-Heaps, J. D. & Northcote, D. H. *J. Cell Sci.* **1**, 109–120 (1966).
- Hepler, P. K. & Newcomb, E. H. *J. ultrastruct. Res.* **19**, 498–513 (1967).
- Hepler, P. K. & Fosket, D. E. *Protoplasma* **72**, 213–236 (1971).
- Hepler, P. K. & Palevitz, B. A. *Rev. Pl. Physiol.* **25**, 309–362 (1974).
- Gunning, B. E. S. & Hardham, A. R. *Endeavour (Oxf.)* **3**, 112–117 (1979).
- Hepler, P. K. in *Cytomorphogenesis in Plants* (ed. Kiermayer, O.) (Springer, New York, 1981).
- Kirchner, M. W. *Int. Rev. Cytol.* **54**, 1–71 (1978).
- Kihlman, B. A. in *Actions of Chemicals on Dividing Cells* (Prentice-Hall, Englewood Cliffs, 1966).
- Dustin, P. in *Microtubules* (Springer, Berlin, 1978).
- Hart, J. W. & Sabnis, D. D. *Planta (Berl.)* **109**, 147–152 (1973).
- Hart, J. W. & Sabnis, D. D. *J. exp. Bot.* **27**, 1353–1360 (1976).
- Rubin, R. W. & Cousins, E. H. *Phytochemistry* **15**, 1837–1839 (1976).
- Slabas, A. R., MacDonald, G. & Lloyd, C. W. *FEBS Lett.* **110**, 77–79 (1980).
- Yadav, N. S. thesis, Michigan State Univ. (1980).
- Mizuno, K., Koyama, M. & Shibaoka, H. *J. Biochem.* **89**, 329–332 (1981).
- Fosket, D. E., Morejohn, L. C. & Westerling, K. E. in *Metabolism and Molecular Activities of Cytokinins* (eds Guern, J. & Péaud-Lenoël, C.) 193–211 (Springer, Berlin, 1981).
- Nash, D. T. & Davies, M. E. *J. exp. Bot.* **23**, 75–91 (1972).
- Davies, M. E. *Planta (Berl.)* **104**, 50–65 (1972).
- Bryan, J., Nagle, B. W. & Doenges, K. H. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3570–3574 (1975).
- Murphy, D. B. & Borisy, G. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2696–2700 (1975).
- Nickell, L. G. & Tulecke, W. *Bot. Gaz.* **120**, 245–250 (1959).
- Nesius, K. K., Uchytel, L. E. & Fletcher, J. S. *Planta (Berl.)* **106**, 173–176 (1972).
- Weisenberg, R. C., Borisy, G. G. & Taylor, E. W. *Biochemistry* **1**, 4466–4479 (1968).
- Studier, F. W. *J. molec. Biol.* **79**, 237–248 (1973).
- Bradford, M. M. *Analyt. Biochem.* **72**, 248–254 (1976).
- Lee, J. C. & Timasheff, S. N. *Biochemistry* **14**, 5183–5187 (1975).
- Schiff, P. B., Fant, J. & Horwitz, S. B. *Nature* **277**, 665–667 (1979).
- Ludueña, R. F., Myles, D. G. & Pfeffer, T. A. *Exp. Cell Res.* **130**, 455–459 (1980).
- Little, M., Ludueña, R. F., Langford, G. M., Asnes, C. F. & Farrell, K. J. *J. molec. Biol.* **149**, 95–107 (1981).
- Schiff, P. B. & Horwitz, S. B. *Biochemistry* **20**, 3247–3252 (1981).
- Parness, J. & Horwitz, S. B. *J. Cell Biol.* **91**, 479–487 (1981).
- Sloboda, R. D., Dentler, W. L. & Rosenbaum, J. L. *Biochemistry* **15**, 4497–4505 (1976).
- Clayton, L., Quinlan, R. A., Roobol, A., Pogson, C. I. & Gull, K. *FEBS Lett.* **115**, 301–305 (1980).
- Kilmartin, J. V. *Biochemistry* **20**, 3629–3633 (1981).
- Roobol, A., Pogson, C. I. & Gull, K. *Exp. Cell Res.* **130**, 203–215 (1980).
- Quinlan, R. A., Roobol, A., Pogson, C. I. & Gull, K. *J. gen. Microbiol.* **122**, 1–6 (1981).
- Cleveland, D. W., Fischer, S. G., Kirchner, M. W. & Laemmli, U. K. *J. biol. Chem.* **252**, 1102–1106 (1977).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).

Genes for immunoglobulin heavy chains and for α_1 -antitrypsin are localized to specific regions of chromosome 14q

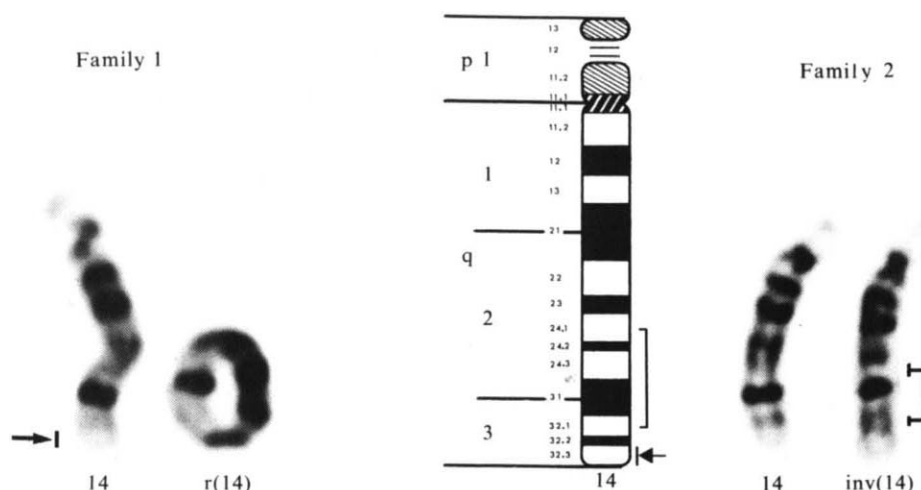
Diane W. Cox, Vera D. Markovic* & Ikuko E. Teshima

The Hospital for Sick Children, 555 University Avenue, Toronto, Canada, and Departments of Paediatrics and Medical Genetics, University of Toronto, Toronto, Ontario M5G 1X8, Canada and

* Genetics Department, Surrey Place Centre, Toronto M5S 2C2, Canada

Studies of somatic cell hybrids have assigned the gene cluster for human immunoglobulin heavy chains μ , γ and α to chromosome 14 (ref. 1). The locus for Gm (GM), a genetic marker on the γ heavy chain, has been shown to be linked to PI, the locus for α_1 -antitrypsin/ α_1 -protease inhibitor (α_1 AT), a major protease inhibitor in human serum^{2,3}. This assignment has now been confirmed by studies of somatic cell hybrids which have assigned α_1 AT also to chromosome 14 (refs 4, 23). From studies of two families having abnormalities involving the long arm of chromosome 14, we have localized GM to the terminal portion of 14q at band q32.3 and PI to a more proximal position, between q24.3 and q32.1. The immunoglobulin genes are within a chromosome region noted for its high frequency of breaks associated with chromosome rearrangement, which occur both spontaneously in cultured lymphocytes and in certain malignancies.

Fig. 1 Chromosomes observed with G-banding. The normal chromosome 14 and r(14) in proband of family 1 is shown with the deleted region marked by an arrow. The normal chromosome 14 and inv(14) in family 2 is shown with breakpoints indicated by brackets. The diagram (centre) shows the prometaphase band pattern according to International Nomenclature²².



Early metaphase chromosomes in cultured lymphocytes were studied by trypsin-Giemsa G-banding⁵. Twenty or more cells of each individual were examined, at least 10 of which were banded. To define the breakpoint of the ring chromosome, C-banding⁶, Giemsa-11⁷ and NOR staining⁸ were also used. PI typing was carried out by isoelectric focusing in polyacrylamide pH 3.5–6, and α_1 AT was quantitated by electroimmunoassay⁹. GM allotypes were determined by haemagglutination inhibition using fresh O Rh⁺ (Rhesus-positive) human erythrocytes sensitized with non-agglutinating anti-Rh antibody¹⁰. All sera were tested at a 1/20 dilution, with antibody for the γ 1 markers G1m (*a*, *x* and *f*) and for the γ 3 markers G3m (*b* and *g*) (reagents were given by Dr Moses Schanfield, American Red Cross). Quantitation of immunoglobulins (IgG, IgM and IgA) was carried out by laser nephelometry (Beckman). Other genetic markers were typed by standard techniques and used to verify the stated family relations: blood groups ABO, Rh and MNS; complement 3 (C3) and properdin factor B (BF) were identified by agarose electrophoresis; and group-specific component (GC) and transferrin (TF) by isoelectric focusing in polyacrylamide gels.

In family 1 (HSC-C4), the proband was an 8-year-old male with mental retardation, seizures, microcephaly and primary pulmonary hypertension. His karyotype, examined from 21 cells the first time and 100 cells on a second occasion, was 46 XY, ring 14 [r(14)]. In the two analyses, 93% of blood lymphocytes had the karyotype 46 XY, r(14), with at least one ring present, while the remainder had one copy of chromosome 14 lacking the ring. Breakpoints were distal to bands p11 and q32.2. The deleted portion of the chromosome is therefore all or part of q32.3 (Fig. 1). Further clinical and cytogenetic details of this patient will be reported elsewhere (V. D. M., R. G. Worton, C. B. Verellen and J. M. Berg, in preparation).

Both parents had normal karyotypes. Results of PI and GM typing, determined on two separate occasions, are shown in Fig. 2. The proband expressed only the paternal GM type, *fb*, and lacked a maternal GM component. This was confirmed in a repeat blood sample from proband and mother; both sera were also tested at 1/2 dilution. This result can be explained if the GM marker (γ -chain locus) is in the deleted region, q32.3. Inherited deletions of γ 1 or γ 3 regions have rarely been reported, however the GM types of I-2 and II-1 (Fig. 2) exclude the possibility of an inherited deletion. Serum concentrations of immunoglobulins are shown in Table 1. The IgG, IgA and IgM concentrations in serum of the proband are normal and do not show a dosage effect. This may, however, be the result of allelic exclusion, by which only one of the alleles for the immunoglobulin chains is expressed in any one immunoglobulin-producing cell, suggesting that the intact IgG region is always expressed when there is a deletion in the homologous chromosome. Genetic types of α_1 AT (PI types) were not informative

as proband and parents were all of PI type M1. Serum concentrations of α_1 AT did not suggest a deletion for this locus; heterozygotes for the null allele of α_1 AT (*PIM*–) have lower concentrations ($59 \pm 6.8\%$ of normal)¹¹ than those found in the proband. Genetic markers were compatible with the stated parentage.

Family 2 (HSC-C3) provided information on localization of the PI locus. The proband was a 7-year-old boy with neutropenia. An altered banding pattern of chromosome 14 was observed in all 26 cultured lymphocytes examined by trypsin-Giemsa G-banding (Fig. 1). We concluded that the altered pattern was due to a paracentric inversion with breakpoints at bands q24.1 and q32.1; karyotype 46 XY, inv(14) (q24.1 q32.1). The inverted chromosome 14 was also found in other relatives (Fig. 2). Genetic markers were compatible with the stated family relations. Results of typing are shown in Fig. 2. If the PI allele, M2 and the inversion are assumed to be coupled in I-2, then there are no recombinants among 11 offspring in this family. These results are compatible with location of the PI locus at the inversion, between bands q24.1 and q32.1. Genes within the inversion would be inherited together because a single cross-over in the inverted region would result in non-viable gametes. As there is an average of 1.76 chiasmata for chromosome 14q (ref. 12), and if a chiasma is equivalent to a cross-over, then a double cross-over, which would be necessary to produce viable gametes, is not likely to occur in the small distance covered by this inversion. Linkage analysis by lod scores supports the location within the inverted region. The inverted region was considered as a marker. Lod scores for the three sibships, two of which are phase known, were obtained from the tables of C. A. B. Smith as given by Emery¹³. Calculations were simplified by assuming I-1 to be of PI type M1M1, where the relative probability for this PI type, based on population frequencies from our laboratory and from PI types of the offspring, is 87%. The maximum likelihood

Table 1 Serum concentration of immunoglobulins and α_1 -antitrypsin in family 1

Protein	Proband	Father	Mother	Controls [mean (95%)] confidence interval
IgG (mg dl ⁻¹)	808	906	998	915 (633–1280*)
IgA (mg dl ⁻¹)	54	134	97	90 (33–202*)
IgM (mg dl ⁻¹)	86	154	186	107 (48–207*)
α_1 AT (% normal)	79	84	100	105 (67–198†)

* Values for children 6–8 y old (from C. R. Jolliff and K. Cost); adult values are higher.

† Obtained from 112 normal adults, in our laboratory. Values for children are similar.

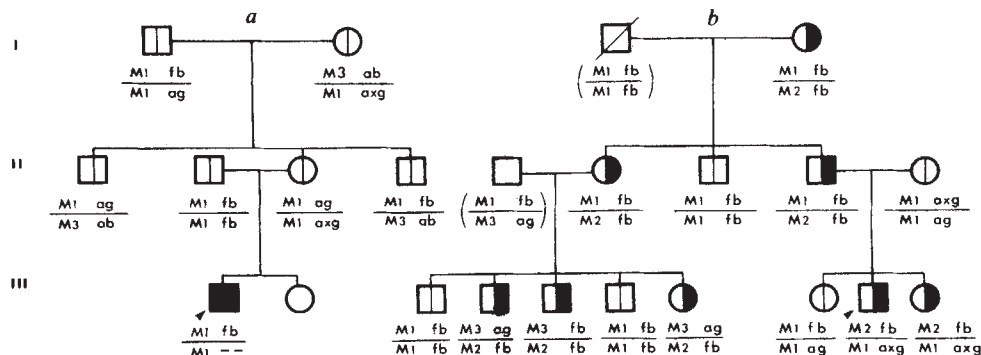


Fig. 2 Pedigrees of family 1 (a) and family 2 (b) showing abnormal chromosomes and the most probable arrangement of PI and GM types. Arrowheads indicate pro-bands. ■ Indicates the presence of r(14); □ indicate the presence of inv(14).

estimate of recombination (B) from plotted values¹² is 0 with a lod score of 3.0. This indicates there is a high probability (1,000:1) of absolute linkage between *PI* and the inversion marker. A map distance of up to 5 centimorgans (cM) (recombination value, $\theta = 0.05$) would not be excluded.

Our studies in family 1 indicate that *GM* is localized at 14q32.3. While we consider the loss of the structural γ -chain locus to be probable, we cannot exclude the possibility that a promoter or sequences necessary for initiation of the γ -chain switch have been lost. Studies of the DNA of the patient carrying the deletion might clarify this question. We have recently learned, however, that *in situ* hybridization studies in the laboratory of P. Leder, using heavy-chain constant-region probes, also confirm that γ -chain loci are localized at 14q32, a region larger than, and including, our region 14q32.3 (I. Kirsch, personal communication).

Data from family 2 indicate that the *PI* locus is in the region q24.1–q32.1. The *GM-PI* linkage in males is 0.26³ for 'non-Z' alleles. As there is an average of 1.76 chiasmata for chromosome 14q (ref. 12), and a recombination value of 0.5 in males is defined as equivalent to 50 cM, the genetic length of 14q is ~88 cM in human males. The recombination value of 0.26 for *GM-PI* falls either at the distal border of q24.3, according to Francke's band patterns based on actual chromosome measurements¹⁴ or at the proximal portion of q24.3 in Fig. 1. However, the distribution of chiasmata in 14q is uneven, with a high concentration in terminal portions of the chromosome, thus the most likely location for *PI* is in a more distal region of the inversion. As it is unknown how the genetic map corresponds to the physical map, we must localize *PI* at q24.3–32.1. Nucleoside phosphorylase, which has been localized at the region q12–13 (ref. 15), should not be linked with *GM*, but may show linkage with *PI*, the minimal distance estimated as being 24 cM.

Our localization of *GM* at band q32.3 may have important biological implications as it is within or close to a region of the chromosome that is prone to breakage. Analysis of spontaneous breaks in cultured lymphocytes has indicated that chromosome 14 was involved in rearrangements more often than any other chromosome¹⁶. The 'hotspots' for breaks were 14q13 and 14q32. In ataxia telangiectasia, an inherited condition in which chromosome breaks occur frequently in cultured cells, the chromosome abnormalities frequently involve the long arm of chromosome 14, sometimes involving a tandem duplication with breaks through q11 and q32 (for review see ref. 17). In Burkitt's lymphoma and in other lymphomas, chromosome breaks are non-random. Burkitt's lymphoma clones frequently contain rearrangements between chromosomes 8 and 14, specifically t(8; 14) (q24; q32)¹⁷. Other abnormalities of chromosome 14 are found with high frequency, consistently with breaks at q32. DNA rearrangements, which occur in the immunoglobulin gene cluster both during the switch from one type of heavy chain to the other and within differentiated cells to produce antibody variability (reviewed in ref. 18), may be

related to this high frequency of breakage. The genes for human light-chain (λ) immunoglobulin chains have been assigned to chromosome 22 and association of this chromosome with translocations in malignancies has been noted¹⁹. The λ gene is also a region which undergoes DNA rearrangement. Malignant clone proliferation may be related to effects on the immunoglobulins, if breaks occur directly within the gene.

A restriction enzyme fragment of human DNA cloned in phage λ CH4-rHs18 has been found to be highly polymorphic, apparently because of DNA rearrangements rather than base pair substitutions or modifications²⁰. This locus, D14S1, has recently been localized at the distal half of 14q²¹. If regions of DNA rearrangement do indeed promote chromosome instability, then it will be of considerable interest to determine whether this locus is the same as, or close to, the immunoglobulin locus.

We thank Dr Joseph Berg who originally made us aware of family 1; Dr Ron Worton for helpful discussions; Drs Moses Schanfield and Linda Walsh for helping to establish GM typing in our laboratory; Patricia Zavitz for obtaining blood samples from family members; and Dr Louis Siminovitch for review of the manuscript. We also acknowledge the technical assistance of Tammy Mansfield, Gail Billingsley, Shirley Smyth, Diane Wills, Vera Foldes, Aileen Brown, Chin Ho, Hannah Ho, Kim Mellick and Kwan Ng. This study was supported in part by a grant from the MRC of Canada (MA5426). I.E.T. was supported by The Hospital for Sick Children, in the laboratory of R. Worton.

Received 7 January; accepted 23 March 1982.

- Croce, C. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 3416–3419 (1979).
- Gedde-Dahl T. Jr, Cook, P. J. L., Fagerhol, M. K. & Pierce, J. A. *Ann. hum. Genet.* **39**, 43–50 (1975).
- Gedde-Dahl, T. Jr, *et al.* *Ann. hum. Genet.* **45**, 143–153 (1981).
- Pearson, S., Tetri, P., George, D. L. & Francke, U. *Am. J. hum. Genet.* **33**, 148A (1981).
- Seabright, M. *Lancet* **ii**, 971–972 (1971).
- Arrighi, F. E. & Hsu, T. C. *Cytogenetics* **10**, 81–86 (1971).
- Bobrow, M., Madan, K. & Pearson, P. L. *Nature new Biol.* **238**, 122–124 (1972).
- Bloom, S. E. & Goodpasture, C. *Hum. Genet.* **34**, 199–206 (1976).
- Cox, D. W. *Am. J. hum. Genet.* **33**, 354–365 (1981).
- Schanfield, M. S., Polesky, H. F. & Sebring, E. B. in *Paternity Testing* (ed. Polesky, H. F.) 45–54 (ASCP-EMS, Chicago, 1975).
- Cox, D. W., Billingsley, G. D. & Smyth, S. in *Proc. 'Electrophoresis 81'*, 505–510 (Verlag-Chemie, Germany).
- Hulten, M. *Hereditas* **76**, 55–78 (1974).
- Emery, A. E. H. in *Methodology in Medical Genetics*, 63–75 (Churchill-Livingstone, Edinburgh, 1976).
- Francke, U. *Cytogenet. Cell Genet.* **31**, 24–32 (1981).
- George, D. L. & Francke, U. *Science* **194**, 851–852 (1976).
- Mattei, M. G., Ayme, S., Mattei, J. F., Aurran, Y. & Giraud, F. *Cytogenet. Cell Genet.* **23**, 95–102 (1979).
- Sandberg, A. in *The Chromosomes in Human Cancer*, 161–165, 400–405 (Elsevier/North Holland, New York, 1980).
- Marx, J. L. *Science* **212**, 1015–1017 (1981).
- Erikson, J., Martinis, J. & Croce, C. M. *Nature* **294**, 173–175 (1981).
- Wyman, A. R. & White, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6754–6758 (1980).
- de Martinville, B., Wyman, A. R., White, R. & Francke, U. *Am. J. hum. Genet.* **34**, (in the press).
- Harnden, D. G. *et al.* (eds) *Cytogenet. Cell Genet.* **31**, 1–23 (1981).
- Darlington, G. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 870–873 (1982).

Isolation and structural organization of the human preproenkephalin gene

Masaharu Noda, Yutaka Teranishi, Hideo Takahashi, Mitsuyoshi Toyosato, Mitsue Notake, Shigetada Nakanishi* & Shosaku Numa

Department of Medical Chemistry and * Institute for Immunology, Kyoto University Faculty of Medicine, Kyoto 606, Japan

Recently, we have elucidated the primary structure of bovine adrenal preproenkephalin by determining the nucleotide sequence of cloned DNA complementary to its mRNA¹. The structure of most of this precursor molecule has also been deduced by Gubler *et al.*² using cDNA sequencing in conjunction with protein sequencing. Bovine preproenkephalin contains four copies of methionine-enkephalin³ (Met-enkephalin) and one copy each of leucine-enkephalin³ (Leu-enkephalin), Met-enkephalin-Arg⁶-Phe⁷ (ref. 4) and Met-enkephalin-Arg⁶-Gly⁷-Leu⁸ (refs 1, 2, 5). The region containing the repeated enkephalin and extended enkephalin sequences, which are each bounded by paired basic amino acid residues, is connected with a cysteine-containing amino-terminal sequence preceded by a signal peptide⁶. We have now studied the relationship between the repetitive structure of preproenkephalin and the structural organization of its gene by cloning a human genomic DNA segment containing the entire gene. We find that the general organization of the preproenkephalin gene is strikingly similar to that of the gene encoding the common precursor of corticotropin (ACTH) and β -lipotropin (β -LPH)⁷⁻⁹ (alternatively designated preproopiomelanocortin), another multi-hormone precursor. Furthermore, the complete mRNA and amino acid sequences of human preproenkephalin have been deduced from the corresponding gene sequence.

A human genomic DNA library¹⁰, kindly provided by Dr T. Maniatis, was screened for phage carrying preproenkephalin gene sequences by hybridization *in situ*¹¹ at 60°C with the bovine cDNA probe¹. The DNA inserts carried by the isolated hybridization-positive phage clones were subjected to restriction mapping and blot hybridization analysis¹² with the cDNA probe (see Fig. 1 legend for details of cloning). Figure 1a shows a restriction map thus obtained for the hybridization-positive 2.0- and 6.7-kilobase pair (kbp) *EcoRI* fragments, which were shown to be adjacent to each other. This restriction map is in agreement with that of the mRNA-coding sequence on human placental cellular DNA constructed by blot hybridization analysis with the bovine cDNA and the 2.0-kbp *EcoRI* fragment as probes, thus assuring that the cloned 2.0- and 6.7-kbp *EcoRI* fragments are derived from a single segment of the human genome. (The *EcoRI* site at the 5' end of the 2.0-kbp fragment is derived from the *EcoRI* linker joining this fragment to the phage vector.) These two *EcoRI* fragments were subcloned in the plasmid pBR322 (ref. 13) for further analysis.

Detailed restriction maps for the mRNA-coding sequences (exons) in the 2.0- and 6.7-kbp *EcoRI* fragments were constructed, and the DNA fragments containing an exonic sequence and its surrounding regions were subjected to nucleotide sequence analysis by the procedure of Maxam and Gilbert¹⁴ according to the strategy indicated (Fig. 1b, c). Comparison of the human DNA sequence determined (Fig. 2) with the nucleotide sequence of the bovine cDNA¹ has enabled us to locate two intervening sequences (introns) in the human preproenkephalin gene. One of them, 469 base pairs (bp) long (intron A), is in the segment encoding the 5'-untranslated region of the mRNA 3 bp upstream from the translational initiation site. The other, ~3.4 kbp long (intron B), interrupts the protein-coding sequence between the triplets encoding amino acids 46 and 47 of preproenkephalin. The segment specifying amino

acids 47–267 and the 317-bp segment corresponding to the 3'-untranslated region of the mRNA are uninterrupted. This assignment of the introns is supported by the general rule for exon–intron boundaries according to which an intron begins with a G-T and ends with an A-G dinucleotide¹⁵. Furthermore, the message strands at the ends of both introns are complementary to the 5'-terminal sequence of U1 small nuclear RNA, as observed for the introns of other genes^{16,17}.

The cloned bovine cDNA used lacked a sequence of ~65 bp encoding the 5'-terminal region of the mRNA¹. To identify the gene sequence of this region and the 5' terminus of the mRNA, we used two methods. First, a 5'-end-labelled DNA probe hybridized to the 5'-untranslated region of the mRNA was treated with S₁ nuclease, and the length of the protected DNA fragment measured by electrophoresing it beside a size marker prepared by chemical cleavage of a sample of the labelled probe¹⁸ (Fig. 3a). Second, we estimated the size of the cDNA transcript formed by reverse transcriptase-mediated elongation of a 5'-end-labelled DNA primer hybridized to a segment of the 5'-untranslated region of the mRNA (Fig. 3b). The S₁ mapping procedure yielded a single DNA fragment which extended to the position corresponding to the C residue on the message strand located 217 nucleotides upstream from the 5' end of intron A (see ref. 19 for this positioning). This position agreed approximately with that of the 5' terminus of the mRNA

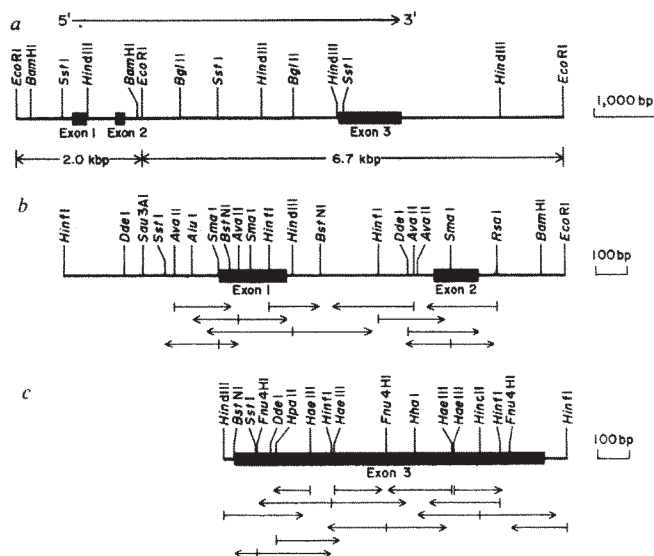
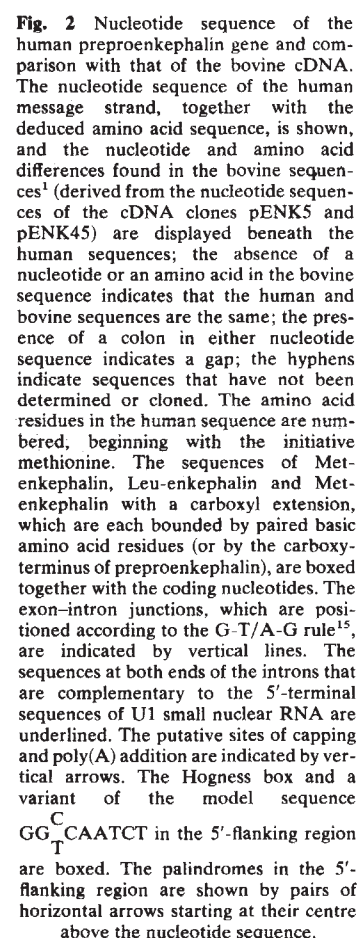


Fig. 1 Restriction mapping of cloned human genomic DNA containing the preproenkephalin gene and sequencing strategy. The library screened was a collection of recombinant phage that carried human fetal liver DNA fragments generated by partial digestion with *HaeIII* and *AluI* and joined to the λ Charon 4A arms with *EcoRI* linkers. The hybridization probe was a mixture of the *PstI*-cleaved inserts of plasmids pENK5 and pENK45 (ref. 1) labelled with ³²P by nick-translation³¹. Five hybridization-positive phage clones were isolated from about 10⁶ plaques by repeated plaque purification. Two of these clones (first type) carried a DNA insert containing six *EcoRI* fragments with approximate lengths of 1.1, 1.2, 1.4, 2.0, 2.9 and 6.7 kbp, and the 2.0- and 6.7-kbp fragments hybridized with the cDNA probe¹². The inserts of the remaining three clones (second type) contained five *EcoRI* fragments approximately 1.1, 1.2, 1.8, 2.9 and 3.2 kbp long, and only the 3.2-kbp fragment was hybridization-positive. The presence of common restriction sites suggested that the two types of clone carried an overlapping DNA segment and that the 3.2-kbp fragment in the clones of the second type constituted a portion of the 6.7-kbp fragment in the clones of the first type. Restriction maps with scales on the right side are given for the DNA segment composed of the 2.0- and 6.7-kbp *EcoRI* fragments (a), a portion of the 2.0-kbp *EcoRI* fragment containing exons 1 and 2 (b) and a portion of the 6.7-kbp *EcoRI* fragment containing exon 3 (c). The direction of transcription is from left to right. For reference, the locations of the exons are shown by closed boxes. All existing sites for the endonucleases indicated are displayed in a; only relevant restriction sites are shown in b and c. The horizontal arrows beneath the restriction maps in b and c indicate the direction and extent of sequence determinations. Further experimental details have been described previously^{8,25}.



On the basis of the results described above, we conclude that the human preproenkephalin gene is ~5.2 kbp long and consists of three exons (designated exons 1, 2 and 3 in the 5' to 3' direction) separated by two introns. If the G-T/A-G rule for exon-intron boundaries¹⁵ is applied and if the redundant A residue at the poly(A) addition site is not counted, exon 1 is composed of 213 bp, exon 2 of 141 bp and exon 3 of 980 bp.

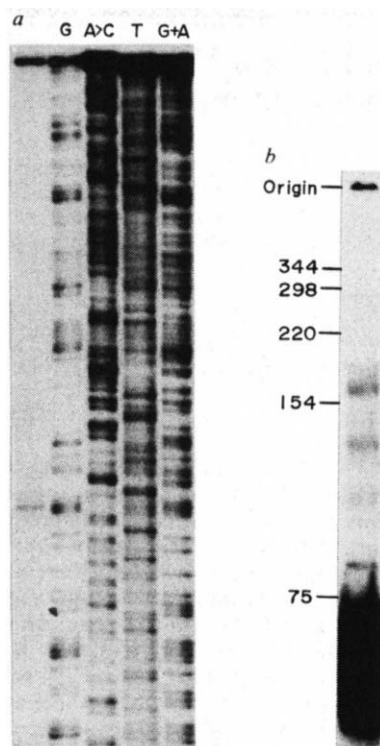


Fig. 3 Identification of the 5' terminus of human preproenkephalin mRNA by S_1 mapping (a) and primer elongation (b). Poly(A)-containing RNA was isolated from human pheochromocytomas obtained by tumour resection according to the procedure used previously for the isolation of rat liver RNA³². The 2.0-kbp *EcoRI* fragment, isolated from the plasmid carrying it, was cleaved by *HinfI*, and the ~650-bp *HinfI*-*HinfI* fragment (located most upstream in Fig. 1b) was isolated and labelled with ^{32}P at the 5' end. In a, this labelled fragment was cleaved by *Sau3AI*, and the ~400-bp *HinfI*-*Sau3AI* fragment (see Fig. 1b) was isolated and used as a probe. The probe (0.03 pmol, 1.2×10^5 c.p.m.) was denatured, hybridized to poly(A) RNA (20 μ g) in 80% formamide at 63 °C for 3 h and digested with S_1 nuclease¹⁸. A sample of the labelled probe was subjected to the G, A+C, T and G+A degradations^{14,33} and used as a size marker. The product of S_1 digestion and the size marker were electrophoresed on 7 M urea/8% polyacrylamide gel. In b, the 5'-end-labelled ~650-bp *HinfI*-*HinfI* fragment was cleaved by *SmaI*, and the resulting products were denatured and electrophoresed on 7 M urea/6% polyacrylamide gel for strand separation¹⁴. The anti-message strand (62 nucleotides) of the smaller *HinfI*-*SmaI* fragment (4 pmol, 1.6×10^7 c.p.m.) (see Fig. 1b) was hybridized to poly(A) RNA (20 μ g) in 0.4 M KCl at 60 °C for 1 h and then subjected to reverse transcriptase reaction³⁴. The cDNA formed was electrophoresed on 7 M urea/6% polyacrylamide gel. The size of the largest cDNA species was estimated to be ~160 nucleotides by comparison with the 5'-end-labelled *HinfI* cleavage products of pBR322 DNA, the sizes of which are given in nucleotides. For further experimental details, see ref. 25.

The structural organization of the preproenkephalin gene, together with the degree of nucleotide sequence homology between the human and bovine protein-coding regions, is schematized in Fig. 4. There are four amino acid additions in the human sequence (Arg-Glu-Asn at positions 85–87 and Asn at position 178). Thus human preproenkephalin consists of 267 amino acids and has a calculated molecular weight of 30,785. The coding regions for the enkephalin and extended enkephalin sequences and the paired basic amino acid residues flanking them exhibit no nucleotide substitution resulting in an amino acid replacement (replacement substitution), except for the substitution of Arg at codon 99 of the human sequence for Lys in the bovine sequence. The coding region for the peptide fragment positioned between the Leu-enkephalin and the preceding Met-enkephalin sequences also shows no replacement substitution. On the other hand, the remaining regions are subjected to some amino acid replacements, although they are fairly well conserved.

We have previously noted that the structural organization of preproenkephalin resembles that of the ACTH- β -LPH precursor in that both proteins are composed of multiple repetitive units (enkephalin or melanotropin sequences) connected with a cysteine-containing amino-terminal sequence preceded by a signal peptide¹. The two proteins are composed of an identical (human) or almost identical number (bovine) of amino acid residues. The present study reveals a striking similarity between the structural organizations of the genes encoding the two proteins. Both genes contain two introns; one of them is inserted in the segment encoding the 5'-untranslated region of the mRNA near the translational initiation site, and the other in the protein-coding sequence near the signal peptide region at an almost equivalent position.

A characteristic of both the preproenkephalin gene and the ACTH- β -LPH precursor gene is that all the repeated enkephalin or melanotropin sequences are encoded by a single large exon (exon 3). When a new gene like an immunoglobulin heavy chain gene was created by duplication of a primordial DNA segment encoding a functional translation unit together with its flanking regions, the organism might have adopted a pre-existing RNA splicing mechanism to eliminate a spacer between the duplicated units²⁴. Unlike the domains of immunoglobulins, the repeated peptide units of the multi-hormone precursors do not function as parts of an integrated protein molecule, but are destined to be separated from one another by proteolytic cleavage to exhibit coordinated biological activities^{1,25,26}. In such a case, the organism might have resorted to a pre-existing protein-processing mechanism to remove the spacer region. The signals specifying the sites of proteolytic cleavage might have existed in the ancestral DNA segment or might have been created *de novo* by drift at the proper positions after duplication. This duplication process could be repeated to produce the characteristic structure of the

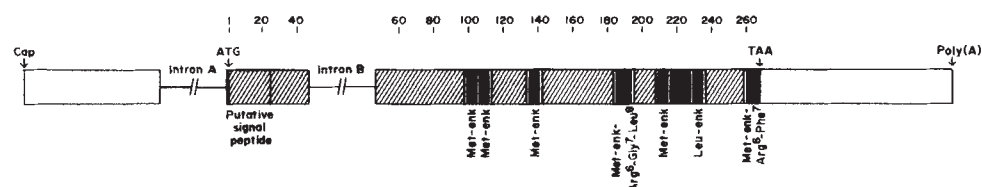


Fig. 4 Structural organization of the human preproenkephalin gene and nucleotide sequence homology between its protein-coding region and the bovine cDNA. The exons are shown by blocks, and the introns by lines; the lengths of the lines do not represent the real lengths of the introns. The sites of capping, translational initiation (ATG), translational termination (TAA) and poly(A) addition are indicated. The segments encoding the 5'- and 3'-untranslated regions of the mRNA are shown by open boxes. Amino acid numbers (see Fig. 2) are given above the gene sequence. The open slits in the protein-coding segment represent the coding nucleotides for the paired basic amino acid residues separating the component peptides, which are displayed beneath the gene sequence. The protein-coding segment is divided into 14 regions bounded by the paired basic residues (or by one of the termini or by the cleavage site of the signal peptide). The regions with and without replacement substitutions are shown by shaded and solid boxes, respectively. The corrected per cent divergence²⁷ calculated for the replacement substitution sites of each region is as follows (codons corresponding to gaps were excluded from the calculation): signal peptide region (amino acid residues 1–24), 8.5; residues 25–97, 10.5; residues 114–133, 13.9; residues 143–183, 16.4; residues 196–207, 9.4; residues 237–258, 5.9. enk, Enkephalin.

multi-hormone precursors. Because such a protein-processing mechanism could overcome the possible deleterious effect of an inexact deletion of an intron, introns which existed originally between the coding regions for the repeated units of the multi-hormone precursors may have been lost, as proposed for the C-peptide-coding region of the rat preproinsulin I gene^{27,28}. Once a flanking peptide, connected with a duplicated unit, has acquired a specific function^{1,25,26}, however, the deletion of an intron would have to be precise, and this would be a rather rare event. In view of the fact that the multi-hormone precursor genes contain, in addition to the coding regions for the repeated units, many segments that apparently represent direct duplications^{1,25}, it seems more likely that the repetitive structure of these genes has evolved by duplication of an ancestral DNA segment at adjacent positions in conjunction with a protein-processing mechanism. On the other hand, the two introns of these genes, located upstream of the region encoding the multiple repetitive units, may have been generated by assembling primordial exons together with their flanking regions during exon shuffling (reviewed in ref. 29). The close resemblance between the structural organizations of the preproenkephalin gene and the ACTH- β -LPH precursor gene suggests that these multi-hormone precursor genes may have evolved by a similar mechanism.

This investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases.

Note added in proof: The cDNA sequence for human preproenkephalin reported recently by Comb *et al.*³⁰ confirms our assignment of the two introns, but indicates that nucleotides 71–157 (numbered beginning with the capping site) can be eliminated by RNA splicing. This spliced mRNA could not be detected by the probe or primer described in Fig. 3 legend, but we have now identified it as a major mRNA species in addition to the mRNA containing nucleotides 71–157 as a minor species. The 5'-terminal 113 nucleotides reported by Comb *et al.*³⁰ can be accounted for by an artefactual inverted repeat of the cDNA sequence starting with the second nucleotide of codon 242.

Received 2 March; accepted 2 April 1982.

1. Noda, M. *et al.* *Nature* **295**, 202–206 (1982).
2. Gubler, U., Seeburg, P., Hoffman, B. J., Gage, L. P. & Udenfriend, S. *Nature* **295**, 206–208 (1982).
3. Hughes, J. *et al.* *Nature* **258**, 577–579 (1975).
4. Stern, A. S. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 6680–6683 (1979).
5. Kilpatrick, D. L., Jones, B. N., Kojima, K. & Udenfriend, S. *Biochem. biophys. Res. Commun.* **103**, 698–705 (1981).
6. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 852–862 (1975).
7. Nakanishi, S. *et al.* *Nature* **287**, 752–755 (1980).
8. Nakanishi, S. *et al.* *Eur. J. Biochem.* **115**, 429–438 (1981).
9. Takahashi, H., Teranishi, Y., Nakanishi, S. & Numa, S. *FEBS Lett.* **135**, 97–102 (1981).
10. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157–1174 (1978).
11. Benton, W. D. & Davis, R. W. *Science* **196**, 180–182 (1977).
12. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
13. Lacy, E., Hardison, R. C., Quon, D. & Maniatis, T. *Cell* **18**, 1273–1283 (1979).
14. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
15. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853–4857 (1978).
16. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283**, 220–224 (1980).
17. Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1877–1879 (1980).
18. Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
19. Sollner-Webb, B. & Reeder, R. H. *Cell* **18**, 485–499 (1979).
20. Weaver, R. F. & Weissmann, C. *Nucleic Acids Res.* **7**, 1175–1193 (1979).
21. Breathnach, R. & Chambon, P. *A. Rev. Biochem.* **50**, 349–383 (1981).
22. Goldberg, M. thesis, Stanford Univ. (1979).
23. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. *Nucleic Acids Res.* **8**, 127–142 (1980).
24. Sakano, H. *et al.* *Nature* **277**, 627–633 (1979).
25. Nakanishi, S. *et al.* *Nature* **278**, 423–427 (1979).
26. Numa, S. & Nakanishi, S. *Trends biochem. Sci.* **6**, 274–277 (1981).
27. Perler, F. *et al.* *Cell* **20**, 555–566 (1980).
28. Bell, G. I. *et al.* *Nature* **284**, 26–32 (1980).
29. Crick, F. *Science* **204**, 264–271 (1979).
30. Comb, M., Seeburg, P. H., Adelman, J., Eiden, L. & Herbert, E. *Nature* **295**, 663–666 (1982).
31. Weinstock, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1299–1303 (1978).
32. Taii, S., Nakanishi, S. & Numa, S. *Eur. J. Biochem.* **93**, 205–212 (1979).
33. Rubin, C. M. & Schmid, C. W. *Nucleic Acids Res.* **8**, 4613–4619 (1980).
34. Houghton, M. *et al.* *Nucleic Acids Res.* **8**, 2885–2894 (1980).

A chicken histone H3 gene contains intervening sequences

James Douglas Engel & Barry J. Sugarman

Department of Biochemistry, Molecular and Cell Biology, Northwestern University, Evanston, Illinois 60201, USA

Jerry B. Dodgson

Departments of Microbiology and Public Health and Biochemistry, Michigan State University, East Lansing, Michigan 48824, USA

The organization of histone gene complexes in higher eukaryotes has long been a focus of study in eukaryotic molecular biology. The best studied and most extensively characterized histone gene families are those of the sea urchin *Strongylocentrotus purpuratus* and *Lytechinus pictus*^{1,2}, which are also the archetypal example of tandemly duplicated genes, each of the four core histone genes and the histone H1 gene being represented once in each tandem repeat unit. Recently, it has become apparent that the organization of histone gene families from higher eukaryotes does not follow these classical examples, a clustered but not tandemly duplicated arrangement having been found in chicken, mouse and humans^{3–5}. In the chicken, the genes are organized in apparently random arrays with no cluster of genes being structurally analogous to another cluster³. In *Xenopus*, there is evidence for some degree of tandem duplication, and yet there is divergence within these clusters⁶. In yeast, the two gene clusters, each containing an H2a and H2b gene, are separate from other histone genes in the organism; furthermore these two clusters are unlinked to one another, but show evidence of close conservation of sequence between the two clusters⁷. During the routine screening and processing of a number of chicken histone gene recombinants we discovered one chicken histone H3 gene which displayed abnormal characteristics with respect to the majority of the genes in our collection. Closer examination of this gene showed it to be structurally different in that it is the first demonstrated example of a histone gene which contains intervening DNA sequences.

Two recombinants initially identified by hybridization to heterologous sea urchin and *Drosophila* histone gene coding sequences are shown in Fig. 1. Figure 1a shows the recombinant insert chromosomal DNA segment of λ CH3d and Fig. 1b shows the position and orientation of an H3 gene contained within λ CH3d and the subclone (pCH3dRI) used for further analysis. Figure 1c shows the sequencing strategy and an abbreviated restriction map in and around the H3 gene contained within λ CH3d, and Fig. 1d–f shows the restriction map, subclone designation and sequencing strategy of the histone H3 gene in recombinant λ CH4a. Mapping and hybridization details of λ CH3d were published previously³.

The histone H3 gene in λ CH3d was chosen for sequence analysis because of strong suggestive evidence from RNA blots that this H3 gene was expressed in embryos (see below). In order to obtain a representative copy of each of the core histone genes for further identification of other, homologous chick histone genes within our isolated recombinants, we chose the H3 gene in λ CH3d at random as an H3 gene-specific sequence. The H3 gene in λ CH4a, however, was chosen for further analysis primarily because of strong hybridization of a mapping blot of this recombinant to cDNA prepared against adult red cell poly(A)⁺ mRNA (data not shown). Since the only chick histone mRNA reported to be polyadenylated is histone H5 (ref. 8), we made the initial assumption that λ CH4a contained a red cell-specific H5 histone gene. Subsequent sequence analysis (see below) proved that this particular locus encodes, rather, a novel histone H3 gene.

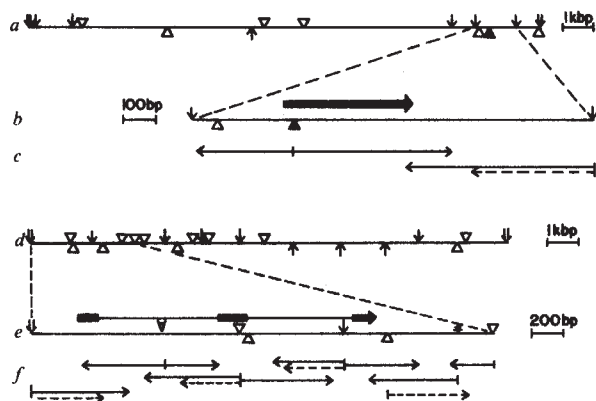


Fig. 1 Chick histone recombinants and DNA sequencing strategy. λ Charon 4A/genomic chicken DNA recombinants¹⁹ λ CH3d (ref. 3) and λ CH4a are depicted in *a* and *d* respectively. Subclones containing the H3 genes from these chromosomal recombinants were prepared by ligation to homologous restriction sites of pBR322. The H3 gene in λ CH3d is contained within subclone pCH3dR1 (*b*; heavy arrow indicates direction and extent of transcription), whereas the coding sequences of the H3 gene in λ CH4a are contained within subclones pCH4a-2.0 (5' and central exons) and pCH4a-2.3 (3' exon). Restriction enzymes used were: $\downarrow$ = *EcoRI*; $\uparrow$ = *BamHI*; ∇ = *HindIII*; Δ = *KpnI*; $\blacktriangle$ = *Sall*; $\ddagger$ = *HinfI*; $\vee$ = *AvaI*. DNA fragments used for sequence analysis were isolated by electroelution²⁰ and labelled at 5' ends²¹ (solid lines, *c* and *f*) or 3' ends²² (dashed lines, *c* and *f*). Chemical reactions²³ and polyacrylamide gel electrophoreses were performed as previously described²³⁻²⁵.

As described previously³, the H3 histone gene in λ CH3d is linked to at least one other chicken histone gene, an H4 gene located in the 5' direction (in transcriptional sense) from the H3 gene of λ CH3d. However, no chicken histone genes other than the H3 gene shown in Fig. 1*e* have been identified on λ CH4a (Fig. 1*d*) or on the chicken DNA sequences of an overlapping recombinant, λ CH5d, which extends approximately 4 kilobases (5' relative to the H3 gene) beyond the end of the λ CH4a insert. However, we cannot yet rule out the possibility that an H1 or H5 histone gene is linked to this H3 gene, or that other histone genes are located in chicken chromosomal DNA close to this gene but beyond the λ CH4a and λ CH5d inserts.

The DNA sequence of the H3 genes in λ CH3d and λ CH4a is presented in Fig. 2. The H3 gene in λ CH3d, like all other histone genes described to date, contains no introns (that is, the DNA sequence indicates that the polypeptide chain is colinear with the H3 polypeptide; Fig. 2, line B). This histone

gene predicts an amino acid sequence which is identical to that of the adult chicken erythrocyte histone H3 polypeptide (line *a*) reported previously⁹. Furthermore, the gene contains 'CCAAT' and 'ATA' polymerase II consensus sequences 5' to the gene (data not shown).

In striking contrast, the histone H3 gene represented by the chromosomal region expanded in Fig. 1*e*, is interrupted by two introns, dividing the coding sequence for this gene into three roughly equal coding parts. This DNA sequence (Fig. 2, line *c*) predicts an amino acid sequence very similar to that of a minor histone H3 polypeptide found in low abundance in somatic avian tissue (for example, in the erythrocyte, liver and perhaps other tissues)¹⁰. However, the predicted amino acid sequence of this gene differs by reciprocal, conservative changes from the chick erythrocyte histone H3.3 polypeptide previously described¹¹ at coding positions 31 and 87 (Fig. 2, line *d*). Thus we cannot be sure whether, in fact, this gene codes for the polypeptide initially described by Urban *et al.*¹¹, or perhaps codes for yet another (low abundance) erythrocyte histone H3 gene.

The DNA sequence of the interrupted gene contains no canonical sequences (outside of the coding sequence) indicative of an expressed gene for poly(A)⁺ mRNAs isolated and characterized to date; that is, within 741 nucleotides 3' to the termination codon we have been unable to find the 'AAUAAA' poly(A) addition sequence found in other eukaryotic poly(A)⁺ mRNAs¹², nor within the 873 nucleotides 5' to the initiation codon do we see evidence for 'CCAAT' or 'ATA' consensus sequences^{13,14} (data not shown). However, the splice junction nucleotides 5'-GT and 3'-AG are present flanking both introns, as have been found in all intron-containing polymerase II-transcribed genes studied to date^{15,16}. 255 bases 3' to the termination codon, one finds an oligo d(A₁₁GA₉) encoded in the mRNA sense strand. This oligo(A) sequence, if transcribed, would account for the selection of the mRNA synthesized from the split gene on oligo d(T) cellulose. We have no independent data to show whether, in addition, a post-transcriptionally added 3' polyadenylate tail also exists.

To ascertain if (and if so, when) the H3 genes were expressed during avian development, RNA blotting experiments were undertaken. The results of this kind of experiment are shown in Fig. 3. The probe was generated from the histone H3 gene subclone of λ CH3d used for sequencing (pCH3dR1; Fig. 1*b*) and hybridizes specifically and with high affinity to RNA prepared from several samples of 4 and 5 day embryonic RNA. Thus, this gene appears to be expressed both in the avian embryo and in embryonic red cells (Fig. 3, lanes 1, 2), but not in late embryonic or mature adult tissues (Fig. 3, lanes 3, 4). When the identical RNA samples were hybridized with similar probes which contained the split gene exons, hybridization was seen to all RNA samples examined (data not shown); however, the pattern of hybridization was more diffuse and always less

Fig. 2 DNA Sequence of linear and interrupted chicken histone H3 genes. The amino acid sequence of the major chicken erythrocyte H3 polypeptide (called H3.2 in ref. 11) is shown in *a*, and is also the sequence predicted by the H3 gene in pCH3dR1 (*b*). Unassigned glutamate/glutamine or aspartate/asparagine residues in the original polypeptide sequence⁹ (between amino acids 73 and 85) are assigned from the DNA sequence (underlined third letter in three letter amino acid code; line *a*). The coding sequence of the split gene in λ CH4a is shown in line *c* above the amino acid sequence for the minor adult H3.3 histone polypeptide¹¹ (identities in the two polypeptides (lines *a* and *d*) are deleted in line *d*). Differences in the two amino acid sequences are underlined; the differences between the prediction from the interrupted gene sequence (line *c*) and the H3.3 polypeptide variant¹¹ are boxed (positions 31 and 87). The position and size of the two introns in the split gene are shown as triangles.



intense than that of the uninterrupted H3 gene. Nonetheless, consistent hybridization was seen to the split H3 gene locus in RNA from every tissue which was examined. We conclude from these data that the histone H3 gene in λ CH3d encodes an RNA species which is primarily embryonic in origin.

Thus the two chick histone H3 genes described here differ strikingly in primary DNA sequence. The H3 gene in λ CH3d has a structure similar to other histone genes reported to date², and hybridizes exclusively to chick mRNAs which are embryonic in origin, whereas the H3 gene in λ CH4a has two interruptions in the coding segment of the gene.

We have not been able to demonstrate the presence of either 5' or 3' consensus sequences common to other eukaryotic polymerase II transcripts in the DNA sequence of the split H3 gene. We are now examining sequences distal to the coding sequence portion of the gene reported here for the presence of these polymerase II consensus sequences. The absence of an observable poly(A) addition sequence (AAUAAA) is not surprising since this gene may not be polyadenylated by a post-transcriptional mechanism, and may appear in the poly(A)⁺ mRNA fraction only because it contains a transcribed A-rich sequence (A₁₁GA₉). The absence of an 'ATA'-type 5'-flanking sequence is in contrast with several other chicken histone gene sequences examined (our unpublished results), but the actual mRNA initiation site may be outside the presently sequenced region of DNA. In particular, it is possible that another intron exists in the 5' transcribed spacer region of the gene, and for this reason we have not yet identified the 'leader' sequence of the mature mRNA in the split gene.

It is interesting from a developmental viewpoint that the histone H3 gene in λ CH3d encodes an embryonic message corresponding to a polypeptide which is likely to be identical to the major adult erythrocyte histone H3 (ref. 9), but nonetheless appears not to be expressed in the adult red cell. This

suggests that a developmental switch occurs in histone biosynthesis in which an identical polypeptide is produced from two separate genes (one embryonic and one adult) which, from the RNA blots (Fig. 3) must differ significantly in nucleic acid sequence. The alternative explanation, that the two genes differ widely in amino acid sequence (for example, in the ambiguous glutamine/glutamate and asparagine/aspartate positions originally unassigned in the amino acid sequence⁹) is less compelling since the histones, as often reported, appear to show extreme amino acid sequence conservation¹⁷.

The structure of the split histone H3 gene in λ CH4a poses several important questions for further investigation. First, is this gene expressed? Several pieces of data suggest that the split gene is transcriptionally active: (1) cDNA prepared to poly(A)⁺ chicken cellular RNA selected by oligo d(T) cellulose hybridizes exclusively and with high affinity to the presumptive exon sequences; (2) the DNA sequence predicts a polypeptide sequence which is very similar to that of an H3 histone found in adult chicken tissue; (3) no internal stop codons are found, in phase, in the three 'exons'; (4) the intervening DNA sequences are bounded by conventional (5'-GT, 3'-AG) splice junctions at the intron-exon borders; (5) in all tissues examined, RNA is transcribed which hybridizes to this locus. We are now examining the strandedness, tissue specificity, abundance and size of the transcription products at this locus. Our preliminary RNA blotting results raise the possibility that the pattern of chicken histone RNA synthesis might well be analogous to *Xenopus* 5S RNA gene expression, in which oocyte-specific 5S RNAs are synthesized only at early times (corresponding to our embryonic H3 gene in λ CH3d), whereas the somatic 5S genes are expressed in both the oocyte and mature frog¹⁸ (corresponding to expression of the split gene).

We thank Drew Rusling for discussion and help with computer analysis and Sheryl Niemiec, Patricia Creatura and Ed Diekhoff for technical assistance. We acknowledge access to the Stanford Molgen Project computer programs and data bank, made available through NIH resources and SUMEX, and helpful discussions thereof with L. Kedes. This work was supported by the NIH (J.D.E.: HL24415 and CA15145; J.B.D.: GM28837), the March of Dimes Birth Defects Foundation (J.B.D.: 1-0759) and the Leukemia Research Foundation; J.D.E. acknowledges the support of an American Cancer Society Faculty Research Award.

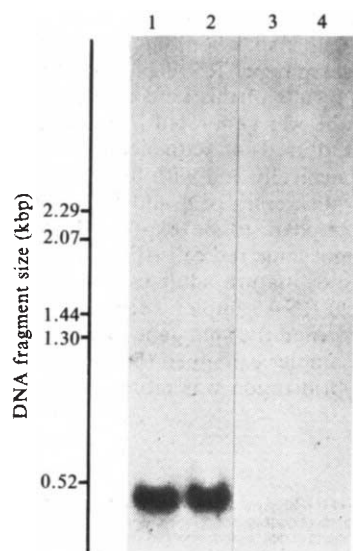


Fig. 3 Analysis of RNAs synthesized by the λ CH3d histone H3 gene. Total cellular RNA and poly(A)⁺ RNA was prepared as described elsewhere²⁶; 5 μ g per lane was reacted with glyoxal in dimethyl sulphoxide, electrophoresed and blotted to nitrocellulose²⁷. RNAs used were: lane 1, 4-day chick embryo total cellular RNA; lane 2, 5-day chick embryo red cell cytoplasmic RNA; lane 3, adult chicken red cell cytoplasmic RNA; lane 4, oligo d(T)-cellulose selected adult chicken red cell cytoplasmic RNA. Hybridization was to nick-translated pCH3dRI (specific activity: 1.5×10^8 c.p.m. per μ g). After hybridization, filters were finally washed in a solution containing total monovalent cation concentration of 20 mM. Thus the homologous T_m should be about 77 °C for a 58% G+C base-pair hybrid²⁸; the filter washes were performed at 60 °C. Exposure time was 9 h.*

Received 28 December 1981; accepted 21 April 1982.

- Kedes, L. H. *Cell* **12**, 321-331 (1978).
- Hentschel, C. C. & Birnstiel, M. L. *Cell* **25**, 301-313 (1981).
- Engel, J. D. & Dodgson, J. B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2856-2860 (1981).
- Sittman, D. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 4078-4082 (1981).
- Heintz, N., Zernik, M. & Roeder, R. G. *Cell* **24**, 661-668 (1981).
- Zernick, M., Heintz, N., Boime, I. & Roeder, R. G. *Cell* **22**, 807-815 (1980).
- Hereford, L., Fahrner, K., Woolford, J., Rosbash, M. & Kaback, D. B. *Cell* **18**, 1261-1271 (1979).
- Molgaard, H. V., Perucho, M. & Ruiz-Carrillo, A. *Nature* **283**, 502-504 (1980).
- Brandt, W. F. & von Holt, C. *Eur. J. Biochem.* **46**, 419-429 (1972).
- Urban, M., Franklin, S., Goldman, P. & Zweidler, A. *J. Cell Biol.* **79**, 119a (1978).
- Urban, M. K., Franklin, S. G. & Zweidler, A. *Biochemistry* **18**, 3952-3960 (1979).
- Proudfoot, N. J. & Brownlee, C. C. *Nature* **263**, 211-216 (1976).
- Elstradatis, A. *et al. Cell* **21**, 653-668 (1980).
- Goldberg, M. L. thesis, Stanford Univ. (1979).
- Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853-4857 (1978).
- Marx, J. L. *Science* **212**, 653 (1981).
- Isenberg, I. A. *Rev. Biochem.* **48**, 159-191 (1979).
- Ford, P. J. & Brown, R. D. *Cell* **8**, 485-493 (1976).
- Dodgson, J. B., Strommer, J. S. & Engel, J. D. *Cell* **17**, 879-887 (1979).
- Girvitz, S. C., Bacchetti, S., Rainbow, A. J. & Graham, F. L. *Analyt. Biochem.* **10**, 492-496 (1980).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Challberg, M. D. & Englund, P. T. *Meth. Enzym.* **65**, 39-43 (1980).
- Smith, D. R. & Calvo, J. M. *Nucleic Acids Res.* **8**, 2255-2273 (1980).
- Dolan, M., Sugarman, B. J., Dodgson, J. B. & Engel, J. D. *Cell* **24**, 669-677 (1981).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
- Schildkraut, C. & Lifson, S. *Biopolymers* **3**, 195-210 (1965).

BOOK REVIEWS

Nuclear power and the people

Eric Ashby

IN January 1977 the Secretary of State for Energy issued a paper entitled *Some Aspects of the Safety of Nuclear Installations in Great Britain*. The importance of this paper lies not in its content but in the reasons for writing it; for it was in part a response to forty-five questions put by the Friends of the Earth. In John Chicken's words, it "established the precedent for questions by interest groups to the Government on nuclear safety to be answered publicly".

Chicken's book recounts the brief history of public concern about hazards from nuclear power and the effects of this concern upon policy. Do not be put off by the elephantine adjectival nouns in the title: the book is a clearly written for the general reader and well documented for the specialist who may want to consult the sources. Its theme is that the terrors of nuclear warfare have, by association, coloured attitudes towards the real, but quite different, risks in the nuclear power industry. The effect of this has been a preoccupation with safety forced upon the Establishment, impulsive reaction — fanned by the media — to every report of an accident and a refusal to accept the evidence that, so far, it has been safer to generate power from uranium than from coal or oil. The climax of the book is an analysis of the Windscale Inquiry published in 1978 and the value of the book is that it puts this Inquiry into historical perspective. It explains (what might otherwise be inexplicable) how nuclear power stations are regarded by some people as polluting in the ancient meaning of the word "pollute": "to render morally impure, desecrate, destroy the sanctity of . . .". As was demonstrated at the Inquiry, the very use of radioactive materials for peaceful purposes is tainted by their possible use for deliberate destruction.

Chicken's sober and modest chronicle begins with a summary of the development of nuclear power and the ways in which it is now controlled. Against this background he sets his story. First, there was the development of a British atomic bomb, described in the superb official history by Margaret Gowing. The decision to make the bomb was taken by a Labour Government under Attlee; or, rather, by a small group of men working in such secrecy that (it is said) even the Cabinet was "excluded from discussion of many of the major decisions on atomic policy". During the whole of the Labour Government's period of office (writes Chicken) "from 1946 to 1951 there was not a single debate

Nuclear Power Hazard Control Policy. By John C. Chicken. Pp.272. Hbk ISBN 0-08-023254-X; pbk ISBN 0-08-023255-8. (Pergamon: 1982.) Hbk £20, \$40; pbk £7.50, \$15.

on Atomic Energy in the House of Commons . . .".

It seemed at first as though the peaceful uses of atomic energy could be developed with as little participation from the public as had occurred in its use for defence. The Atomic Energy Authority was set up in 1954. Its first report evidently did not anticipate hostility from the public. It carried a mild reassurance about risk: the reactors would "present no more danger to people living nearby than many existing industrial works that are sited within built-up areas". In 1957 Britain's first nuclear power station was opened and it was announced that more would follow; a Safety Branch of the Authority was created: all without any serious signs of concern among the public. It was a period of "quiet acceptance"; policy for energy was being left to the experts.

In the 1960s this quiet acceptance was overturned by a worldwide surge of disenchantment with technocrats and the establishments that employed them. The episode is too close to us for an objective historical analysis to be made of it. Chicken concentrates upon its impact on policies for energy, though he quite rightly sees that nuclear power was only one of several issues used by the so-called "counter culture" as tools for protest. The movement began in the United States before it reached Britain, fuelled by resentment about the Vietnam war and the denial of civil rights to minority groups. It spread to an indictment of industry, symbolized by Ralph Nader's withering criticism of negligence over safety in the design of automobiles. A wave of anxiety about pollution and a mood of guilt about exploitation of the environment swept across the affluent industrial nations. Radiation is the most ominous form of pollution; its military uses, its temptation for terrorists, its association with cancer, its persistence if released into the environment, the delayed action of its alarming toxicity, even its invisibility: all combined to make nuclear power a prime target for suspicion and protest.

Anti-nuclear lobbies work at two quite different levels, often simultaneously. At one level they question the need for nuclear power at all; they want an energy policy based on coal or, better, renewable

resources, and much greater economy in the use of energy. They are willing to accept a lower material standard of living to achieve these ends. At the other level are those who question the safety of specific reactors on specific sites. It is at this second level that assurances about hazard-control become essential if there is to be a viable nuclear power industry. In 1975 two events occurred that concentrated attention upon hazard-control: one was the accident to a reactor at Browns Ferry in Alabama which — though nobody was hurt — was useful ammunition for the anti-nuclear lobbies. The other event was the publication of the Rasmussen Report, which became — and remains — the most useful weapon in the armoury of pro-nuclear lobbies. At the time of the Presidential Election in the United States in 1976 the call for a moratorium in the building of nuclear reactors became an election issue. In the following year tens of thousands of demonstrators challenged the building of a fast reactor in France. In Britain there was comparative quiet on these issues until British Nuclear Fuels proposed to extend their fuel reprocessing plant at Windscale. The Government wisely decided to throw the whole issue open to public scrutiny and discussion.

There is already a shelf-load of books, articles and pamphlets about the Windscale Inquiry. Nevertheless Chicken's careful analysis of the decisions about risk and the influence of public interest groups on these decisions is a valuable addition to this literature. The important and encouraging point he makes is that the public interest groups, despite their reliance on voluntary subscriptions to pay the costs of conducting their case, did have a significant influence on the findings of the Inquiry and taught the proponents of the proposal lessons they would not have learnt without the dedicated and competent efforts of such bodies as the Friends of the Earth.

Chicken's analysis leads him to make some tentative suggestions about procedures for involving the public in decisions about what they regard as "acceptable risks". Despite criticisms (some of which, for example in *The Ecologist*, quoted on p. 172, were contemptible) the style of the Inquiry was satisfactory. But there is room for improvement in three respects. First, it is important that public interest groups should be financed from public funds for the significant part they play in such inquiries. Chicken does not mention the

successful way in which this has been done for similar inquiries in Canada. Second, it is essential that the public interest groups should have access to all the technical data possessed by the proponents. Third, and equally essential, Parliament, as the legally constituted forum for the ventilation of public opinion, should participate in decisions as to how safe is safe enough.

Finally, a word about the style and production of the book. John Chicken is to be congratulated on the impeccable way in which he refers the reader back and forth to the details of his complex chronicle: the book is admirably signposted. As to production, the book has been published from camera-ready copy which is a good thing because it keeps the price at a reasonable level; but there are a lot of misprints: surely these could have been corrected as easily in typescript as in ordinary set type. □

Eric Ashby is a Fellow of Clare College, Cambridge. He is joint author with Mary Anderson of The Politics of Clean Air published by the Clarendon Press late last year.

Journals' review issue 1982

On October 7th *Nature* will publish a second review supplement devoted to science journals. Last year's supplement, covering journals first published between January 1978 and May 1980, appeared in *Nature* 293, 341-369; see p.343 of that issue for details.

Criteria for inclusion of a journal in the 1982 issue are that:

- the first number appeared, or the journal was re-titled, between June 1980 and May 1981;
- it is published at least three times a year;
- the main language used is English.

Broadly, periodicals of professional interest to scientists will be considered for review, with the exception of abstracts' journals.

In addition it is hoped to cover publicly available newsletters, first published between January 1978 and May 1981, in that issue.

Publishers and societies are invited to submit four sample issues of periodicals satisfying the above criteria, *including the first and most recent numbers*, to the Review Editor, *Nature*, 4 Little Essex St, London WC2R 3LF, England (London 836 6633 ext 2570) as soon as possible.

● **Oliphant** by Stewart Cockburn and David Ellyard (reviewed in *Nature* 296, 685; 1982) is available in the UK from W. Heffer, 20 Trinity St, Cambridge CB2 3NG or C.M. Watkins, 44 Madrid Rd, London SW13. Price is £14.90 + post and packaging.

● **Vernon Booth's Writing a Scientific Paper and Speaking at Scientific Meetings, 5th Edn** (reviewed in *Nature* 296, 499; 1982) can be obtained from The Biochemical Society Book Depot, PO Box 32, Commerce Way, Colchester CO2 8HP, UK.

Catastrophes in science and engineering

Colin Upstill

Catastrophe Theory for Scientists and Engineers. By R. Gilmore. Pp.680. ISBN 0-471-05064-4. (Wiley: 1981.) £32.15, \$61.
Instabilities and Catastrophes in Science and Engineering. By J.M.T. Thompson. Pp.226. Hbk ISBN 0-471-09973-2; pbk ISBN 0-471-10071-4. (Wiley: 1982.) Hbk £14.50, \$34.75; pbk £8.25, \$19.50.

THE term "catastrophe theory" is open to wide interpretation, depending upon the scope accorded it. René Thom considered catastrophe theory to be a theory of general morphology, a topological bifurcation theory for systems and bifurcations of very general type, and this usage is now common. Under this broad roof lives a great body of work including such topics as bifurcation theory, dynamical systems theory, singularity theory and elementary catastrophe theory. The lines of demarcation between topics are neither fixed nor of zero width: something of a consensus is emerging, but many authors have used, do use, and no doubt will use, these and other terms without clearly stating what they mean by them.

Elementary catastrophe theory (ECT), the study of the equilibria of gradient systems, is the part of catastrophe theory which embodies the most simplifying assumptions, and is the most developed and accessible part of the theory. The qualifying adjective is often dropped for ease of exposition; thus catastrophe theory may mean something very general or something quite specific. When, as sometimes happens, the same author uses the term in both senses, the potential for confusion is maximized. Neither Gilmore nor Thompson is actually guilty of this, but there is a marked difference in the manner in which they deal with the welter of jargon.

In the first three pages of an excellent text, Gilmore explains briefly and clearly what catastrophe theory is, and what dynamical and gradient systems are and how they and ECT fit together. In the first of four parts of the text, he goes on to "present Elementary Catastrophe Theory from an elementary point of view" in a way which is well tailored to the needs and background of most physical scientists and engineers, defining technical terms crisply and clearly without recourse to the cryptic notations and abstraction of the mathematical literature.

Thompson's book is not so much a text as a review of many (but by no means all) current applications of catastrophe theory in the physical sciences and in engineering. It performs this task splendidly, provided that one has some knowledge of the subject to start with. For, despite the claim of the preface that the presentation is "at the level of popular magazines such as *Scientific American* or *New Scientist*", a great

variety of jargon, both catastrophe theoretic and specific to the specialist fields covered, is used without ever being explicitly defined. The first use of many such terms is flagged by their being italicized, but setting something in italics is no substitute for explaining what it means. In collaboration with G.W. Hunt, Thompson is preparing a companion book to deal with much of the underlying theory, so it may well be that the decision to use unexplained technical terms is a deliberate one. Nonetheless, I believe that the book is considerably less accessible to many of the intended readers than it would have been with the occasional extra sentence of definition and explanation. This problem is exacerbated by Thompson's use of a transcript of his inaugural lecture as a basis for the introductory chapter, which gives a brief historical survey of the subject and an overview of the contents of the rest of the book. This is excellent for the *cognoscente*, but at the price of presenting the reader new to the subject with a high density of unexplained technical terms, and with a number of complicated diagrams referred to by phrases such as "the familiar such-and-such", which will almost certainly be anything but familiar to him.

Having taken Thompson to task for not writing the book the cover and the preface lead one to expect, let me repeat that the book he has actually written is a good one, well illustrated — essential when dealing with so many intrinsically geometric notions — and with copious references. He distinguishes clearly between static and dynamic bifurcations whilst showing how they are related, particularly in the context of engineering structures; among the other instability phenomena reviewed are gravitational stellar collapse, atomic lattices, biochemical reactions (including the celebrated Brusselator), population dynamics and ecology, hydrodynamical turbulence and the chaotic dynamics of strange attractors.

In Part 2 of his text, Gilmore considers some applications of ECT, and in Part 3 he first takes us beyond ECT into dynamical systems theory and then considers some examples. These parts of his text are, broadly speaking, complementary to Thompson's book; there is some overlap, but most of the topics Gilmore discusses are not considered at any length by Thompson, and vice versa. Gilmore deals with such topics as phase transitions in classical and quantum mechanical systems, climatology, and optical caustics and diffraction patterns. His discussion of optics is noticeably out of date in comparison with that of other topics, but he neither makes nor implies any claim to be reviewing the state of the art, giving few references and more relaxed and yet deeper discussions than Thompson.

Indeed, Gilmore's relaxed style will not be to everyone's taste; for example, after a page or two describing a standard treatment of some phenomenon, he concludes with the acerbic aside "In short, an intractable problem is replaced by an intractable problem". It is a style he handles well; it could so easily be overdone, but in my view it is not, and greatly enhances readability and thence understanding. The most formal section, also the shortest, is Part 4, in which he presents in some detail the mathematics underlying ECT. In contrast with Part 1, this is (necessarily) couched in sophisticated mathematical language. It would have been easy to have included this material in Part 1, and many readers would not then get beyond it: it is greatly to the benefit of the book that this mistake was avoided. Coincidentally, Part 4 is the least

satisfactory part of the book, when judged against the standard work in the field, the classic text of Poston and Stewart (*Catastrophe Theory and its Applications*; Pitman, 1978) whose treatment of Thom's theorem, transversality, determinacy and unfoldings stands head and shoulders above Gilmore's. That criticism is more than offset by the fact that, particularly for the reader not well versed in the language of modern pure mathematics, the first nine-tenths of Gilmore's text is excellent.

In his epilogue, Gilmore asks the question "Is catastrophe theory useful and important?" and concludes by asserting his belief that it is. I share that view, and both his and Thompson's books are convincing evidence in support of it. □

Colin Upstill is at the H.H. Wills Physics Laboratory of the University of Bristol.

Islands or East Africa. The problem is made more difficult by the way in which Gill attempts to distinguish tholeiitic and calc-alkaline andesites. While he recognizes the importance of iron content, he ignores the more reliable compositional criteria of Irvine and Baragar and makes no mention of the high alumina and normative plagioclase contents that many petrologists consider compositionally diagnostic and crucial to any petrogenetic interpretation of calc-alkaline andesites.

Because much of the evidence bearing on these rocks comes from geochemical relations, any book of this kind must deal with complex trace-element patterns and isotopic ratios, and with how these may be related to the processes of magmatic evolution. Gill uses these data to weigh the relative importance of melting subducted oceanic lithosphere, assimilating crustal material and differentiating basaltic magma. In this way he systematically assesses essentially all the current ideas on the origin of andesitic magmas. While he does not hesitate to give his own judgement on these questions, he does so in the context of an overall appraisal of the evidence, so that readers can accept or reject Gill's conclusions according to their own view of the facts.

No two persons would agree as to how much weight should be given to various types of evidence bearing on the origins of these rocks. I would personally have preferred more data on geologic relations and less on trace-element geochemistry. It astonishes me that anyone can write an entire book about andesites and never present a single geologic map or structural section of an andesitic volcano. The book, like most modern work which covers orogenic igneous rocks, deals almost exclusively with the structural relations observed today around the Circum-Pacific and gives much less attention to the geologic past. If it is important to note that all active andesitic volcanoes are in well-defined belts close to convergent plate boundaries, is it not equally important to stress that the same has not always been true in the past? Many rules of thumb based on modern relations fail when we look beyond the Quaternary, and I suspect that one can learn more about andesites from the anomalous cases that violate our conventional models than from all the examples that neatly fit them.

These, of course, are petty differences of view, and they detract in no way from my unreserved recommendation of this book as a timely, scholarly survey of a very complex subject. The sincerest praise I can offer any book is to say, as I can of this one, that I have adopted it as a text in my graduate-level petrology course. No student of orogenic igneous rocks should neglect adding it to his professional library.

Alexander R. McBirney is a Professor in the Department of Geology at the University of Oregon, Eugene.

All around andesites, and beyond

Alexander R. McBirney

Orogenic Andesites and Plate Tectonics. By J.B. Gill. Pp.390. ISBN 3-540-10666-9. (Springer-Verlag: 1981.) DM 98, \$44.60.

A THOROUGH, up-to-date review of orogenic andesites has long been needed and is especially welcome at this time, when many petrologists are taking stock of this group of rocks and their interpretation in terms of accepted concepts of plate tectonics. In undertaking this formidable task, James Gill has performed a notable service to petrology. His book is certainly the most comprehensive compilation of data and ideas that has yet been attempted.

The book summarizes almost every aspect of the occurrence, composition and genetic relations of andesites. It reviews their structural and tectonic setting, rheological properties, eruptive behaviour, geochemical and mineralogical features, compositional variations in time and space, and all the major petrogenetic theories proposed to explain their relations to plate tectonics. This mass of data is presented in a concise, well-organized fashion. In his evaluation of the various schemes which have been proposed to account for orogenic igneous rocks, Gill turns much of his attention to their relationship to subduction, and a large part of the book is an attempt to reconcile the occurrence and unusual features of the rocks with a party-line view of plate tectonics. In doing so, he has made a conscientious effort to mention all observed relationships, together with every known exception; the result is a text which tends towards the encyclopaedic, with awkward sentences that one must read two or three times to sort out an elusive trivial point. The style improves, however, when

Gill comes to a question of interpretation that clearly interests him. The text then blossoms and becomes quite readable.

The book is burdened by the long-standing difficulty of defining what rocks should be included under the name andesite. The reader encounters this problem in the very first sentence, which tells him that "active volcanoes on Earth erupt andesite more than any other rock type". Before recovering from this, he reads, further down the same page, that these andesitic volcanoes include Kilimanjaro and Hekla! The explanation for these startling assertions soon becomes apparent when Gill gives his definition of andesite as any hypersthene-normative rock with a silica content between 53 and 65 weight per cent. Although this definition is offered in the interests of "simplicity" the resulting confusion in the mind of the reader persists throughout the book, for one is often in doubt as to whether, in referring to "andesites", Gill means calc-alkaline andesites or some other intermediate sub-alkaline rock. Much needless discussion is devoted to rocks that have no relevance to andesitic volcanism as most geologists perceive it.

There is a certain logic in including among andesites the important group of intermediate tholeiitic rocks of island arcs that are clearly related to subduction and share many genetic features in common with the "true" calc-alkaline andesites with which they are closely associated, sometimes even in the same volcano. It seems illogical not to refer to these rocks as tholeiitic andesites, but most petrologists would certainly agree that they are quite distinct from rocks of similar composition in settings such as Iceland, the Galapagos

How ionizing radiation affects cells

R.B. Setlow

The Molecular Theory of Radiation Biology. By K.H. Chadwick and H.P. Leenhouts. Pp.377. ISBN 3-540-10297-3/0-387-10297-3. (Springer-Verlag: 1981.) DM 128, \$67.20.

SOME of the best epidemiological data relating carcinogenic risks to environmental hazards comes from research into the consequences of exposure to ionizing radiation. The recent recalculation of dosimetry at Hiroshima and Nagasaki has just about eliminated any effects in these populations ascribable to neutrons, and, because gamma ray data are not robust below 30 rads, there are, for practical purposes, no human data at low doses or for particles of higher linear energy transfer. Thus, even in this best case, one needs models and theories for extrapolating from acute high dose effects to chronic low dose ones. If such a problem cannot be solved for ionizing radiation with its relatively good dosimetry, the chances are negligible that it will be solved for populations exposed to chemical agents.

In this book, Chadwick and Leenhouts describe the development of a theory that can be used to extrapolate the biological effects of ionizing radiation to low doses, many different LET particles, and to interactions between radiation and chemical agents for end points such as cytotoxicity, mutation, chromosome aberrations and carcinogenesis. Their theory is based on the premise that DNA is the target and that double-strand breaks in it are the molecular lesions responsible for the biological effects of radiation. The authors are led to this assumption for a number of reasons, the principal one of which is that the dependence of many biological effects on dose contains both a linear and a square term. The linear term is thought of as representing double-strand breaks arising from the passage of a single ionizing particle close to DNA, and the square term represents double-strand breaks arising from two nearby single-strand breaks made by two ionizing particles.

The text is a careful and complete exposition of the implications of these assumptions. Of necessity, other concepts and hence additional parameters and suppositions must be introduced, such as allowance for the repair of double-strand breaks and the assumption that *all* radiation mutations arise from double-strand breaks. Thus, the analysis rapidly accumulates other variables that are often difficult to evaluate independently and that must be adjusted to fit the experimental results. The general problem faced by the authors is equivalent to that of determining enzyme mechanisms at the molecular level from enzyme kinetics: it cannot be done. Certainly, there is no unique solution to the problem especially because there have been

relatively few good determinations of double-strand breaks as a function of dose and, because of a lack of basic knowledge of events at the molecular level, of how ionizing radiation affects even simple end points — such as the inhibition of the initiation of clusters of eukaryotic replicons — and inhibits the progression of cells through the cell cycle.

The authors do themselves an injustice by overstating their case. The assumption made early in the book that double-strand breaks are important becomes a fact later on. Such confidence is misplaced because it ignores data on the effects of radiation on the replicative form of $\Phi\chi 174$, in either protective or dilute solution, in which inactivation is not correlated with double-strand breaks but with some type of base damage. Moreover, although fibroblasts from individuals with the disease ataxia telangiectasia are more sensitive to the cytotoxic effects of ionizing radiation than are normal cells, and show more chromosomal aberrations as a result of radiation, they are hypomutable. Hence, not all of the biological end points are comparable and the simple theory expounded in the text seems to need major revision based on solid biology.

These are obviously serious criticisms. Nevertheless *The Theory of Radiation Biology* is an important book for radiation and environmental biologists for, if nothing else, it is sure to stimulate new experiments and ideas. □

R.B. Setlow is a Senior Biophysicist at the Brookhaven National Laboratory and a member of the Board on the Effects of Radiation of the US National Academy of Sciences.

Counting calories

John R. Krebs

Physiological Ecology: An Evolutionary Approach to Resource Use. Edited by Colin R. Townsend and Peter Calow. Pp.393. Hbk ISBN 0-632-00555-6; pbk ISBN 0-632-00617-X. (Blackwell Scientific/Sinauer: 1981.) Hbk £21, \$38; pbk £11.80, \$23.60.

I WAS once told by a fellow graduate student, trained in herpetology, that physiological ecology was the science of shoving thermometers up lizards' cloacas. A quick check in the subject index of this volume revealed that, given the veracity of my colleague's assertion, things have come a long way in the past 15 years.

The book, comprising a collection of original articles and intended as a text for

undergraduate courses, is mainly about how plants and animals get their daily calories and what they then do with them. The unifying theme linking such diverse topics as photosynthesis, digestion, cellular repair and social behaviour is that all aspects of earning and spending calories can be treated from a functional viewpoint as optimization problems. Different chapters illustrate different degrees of progress in applying this approach. In some areas, for example leaf design and foraging behaviour, the first generation of explicit models has been formulated and tested fairly thoroughly, while in others such as defence mechanisms and social behaviour the questions have, in the main, been formulated only qualitatively.

The articles also vary in the extent to which they present totally new ideas. R.M. Sibly, in one of the more original chapters, provides a fresh and stimulating approach to digestive strategies. He starts with a version of the marginal value theorem to predict optimal retention times for food in the gut, and leads via a discussion of gut morphology to consider why food empties from the stomach at an exponentially decreasing rate from the time of the last meal (no simple answer to this question emerges, but the relevant data and hypotheses are neatly brought together). The highlight of the book in terms of entertainment value is D.H. Janzen's fireside chat about defences against predators and parasites. Among the aphorisms that abound in this chapter is the answer to "Why should some species of seed be so well defended when the plant makes so many?" — "... for the same reason that a young man protested at being sent to Vietnam". At the same time, I should add that Janzen does not, to my mind, succeed in pushing the study of plant and animal defences from a series of intriguing evolutionary stories to forming a coherent theory. He appeals for a change in attitude from "descriptions of yet another alkaloid [to measuring] fitness gains and costs that the organism accrues with and without its defences" without stating explicitly how these hard-won measurements will be used to make a general theory of defences.

Some contributors — for example Pianka on resource acquisition and allocation among animals, and Gosling and Petrie on the economics of social organization — try to cover too wide a field, and in so doing do not analyse any single problem beyond introductory textbook level, while others — Kirkwood on repair and its evolution, for example — deal in detail with a more restricted range of topics. Overall, however, the book is a worthy collection of reviews which covers a range of subjects that are not normally assembled in one volume. □

John R. Krebs is a Nuffield Foundation Science Fellow at the Edward Grey Institute of Field Ornithology and E.P. Abraham Fellow of Pembroke College, Oxford.

The principles of vision enumerated

Oliver Braddick

A Taxonomy of Visual Processes. By William R. Uttal. Pp.1,097. ISBN 0-89859-075-2. (Lawrence Erlbaum: 1981.) \$100, £50.

THE stated aim of Uttal's book is to "... systematize the data base of perceptual psychology into a comprehensive intellectual scheme...". When the route to this goal runs through 1,100 pages and over 3,000 references, it must involve to a fair degree the exposition of the "data base" itself, and there is here a generous view of its extent, including the discovery of alpha particles, how a cathode ray tube works, the embryology of frog rods and the theory of rainbows. More conventionally, the field of vision is covered from photometry to geometrical illusions. It would not be reasonable to expect a single author to compete in authority over this range with the several multi-author handbooks of similar scope that have appeared in recent years, and Uttal does not do so.

The author shows a propensity for organizing his data base in lists, which rather blunt the reader's appreciation of what might be the relative significance of the information. This culminates in the closing pages in a list of 67 "general principles" from which we may sample No.4:

For millennia, philosophers, scientists, and laymen alike have felt a deep-seated urge to explain the phenomena of perception. . . . No perceptual theory, however, has ever withstood the test of time . . .

and No.59: "The magnitude of many illusions depends on the orientation of the pattern . . .". Several of the lists contain direct contradictions that remain unresolved and without comment by Uttal. The general air is one of abstracting rather than digesting the scientific literature; opportunities for integration are conspicuously missed, for example the use of the temporal modulation transfer function to express data on the visual response to a variety of time-varying stimuli.

But it would be unfair to dwell on Uttal's presentation of the data base, rather than his intention of providing a framework for systematizing it, the "taxonomy" of the title. This taxonomy is one of sequential levels of processing, from Level 0 in the optics of the environment and the eye, to Level 5 which consists of active, attentive manipulations of perceived information which are reserved for another projected book. Most visual scientists use some such idea of levels, albeit informally, as a way of compartmentalizing problems. This is not to say that it is the best principle for the organized presentation of our knowledge of visual perception, and several disadvantages become apparent in Uttal's attempt to do so. One is that there are many phenomena about which we know a good deal, and which are clearly important, but which

we cannot partition among levels with any certainty. Ironically, Uttal's own experimental work largely falls into this category, and into a chapter on temporal interactions which he has to present as a frank digression, unassignable to any of his levels on present understanding. In other cases, the organization of the book by levels makes it difficult to evaluate alternative explanations critically. For example, the issue of lateral inhibition versus more global processes in brightness contrast is raised, with vigorous assertions of the importance of the latter, but the plan of the book does not allow them to confront each other. Another consequence is an absence of any functional perspective. The purposes for which vision is used cut right across these levels, so they get little attention here.

A reason for the taxonomic approach becomes apparent in a chapter entitled "Mezzolog: On the Limits of Neuro-reductionism". Uttal is concerned that much current theorizing, providing speculative neural bases for perceptual phenomena, is premature. Once again, this view is expressed centrally as a list: "Questionable Dogmas", Nos 1-18. Some of the dogmas are general suppositions about the relation between brain activity and perception, and indeed some of these are quite widely believed, at least implicitly. Others (for example No. 15: "Metacontrast results from peripheral lateral inhibitory interactions that diminish the strength of the conducted neural signals") may be specific instances of neuroreductionism, but are not necessarily widely accepted even by those committed to neural reductionist explanations of the phenomenon concerned. Dogmas of these two kinds need rather different kinds of critique, but it is not clear that the difference is appreciated.

The role of the taxonomy is to draw a sharp line between Levels 1-2, where Uttal offers accounts in terms of interaction of neural signals, and Level 3-4 where the processes are "interpretive", the neural representations involved are not at all understood, and theories are psychological or purely formal.

Uttal's charge of premature reductionism is just in many cases. However the structure and style of his book do not favour a critical marshalling of the arguments on the neural basis of

specific visual phenomena. Indeed, given the scale of the book, any particular issue is pursued in surprisingly little depth. The temptation of encyclopaedic breadth seems to have deflected the author's purpose of exploring the limits of current neural explanation, which is a pity, for that purpose is a very worthwhile one.

Uttal recounts in his preface that when he "chose to become an explicit generalist rather than a totally dedicated . . . laboratory scientist" the reaction of some of his colleagues was such that "some of the experiences encountered were non-supportive to say the least". He defends his role as one of "analysis, criticism, integration and synthesis". I fear that generalism has caused these activities to be diluted and dispersed to a rather unsatisfying degree. □

Oliver Braddick is a Lecturer in Experimental Psychology at the University of Cambridge.



The art of M.C. Escher: this picture, Möbius Strip II (Red Ants), is reproduced from *Escher*, edited by J.L. Locher and recently published by Thames & Hudson, London, and Abrams, New York. The book, containing both a copiously illustrated catalogue of Escher's graphic work and a biography, costs £35, \$65.

Inner machinery of the silicon chip

Andrew Holmes-Siedle

Instabilities in MOS Devices. By J.R. Davis. Pp.190. ISBN 0-677-05590-9. (Gordon & Breach: 1982.) \$24.50.

MOST silicon integrated circuits are smaller in size than a fingernail and contain many thousands of working elements, the most common of which is the metal-oxide-semiconductor (MOS) transistor. In this, all of the important electronic activity takes place in a minute volume, a layered structure composed of crystalline silicon, thermally-grown silicon dioxide and a conductor. Attempts are always being made to reduce the unit size of the transistor element, but during such development it is crucial to maintain the performance of the intimately interwoven materials of the layered structure. *Instabilities in MOS Devices* is a discussion of the physical processes occurring in this metal-oxide-silicon system.

In MOS structures under electrical fields, electrons, holes and ions sometimes migrate. This movement produces the instabilities in electrical properties referred to in the title. In their constant endeavour to ensure that instability effects do not wreck new designs, physicists regard the silicon dioxide layer as the key element; it is a complex film but, in this work, can often be regarded as a simple, wide-band-gap insulator region, sandwiched between the "electron sea" of a metallic electrode and the "bendable" energy bands of the semiconductor. In this context, the instabilities are interpreted either as the unwanted carrying of holes and electrons across the two high, interfacial potential-energy barriers by tunnelling or carrier excitation; or the internal drift of various ionic species, often sodium or hydrogen. The transport medium can often be regarded as very pure fused silica, which allows us to use the large fund of knowledge of the physics of crystalline and fused quartz — makers of science policy might note that this is a good example of the benefits imparted to technology by apparently unrelated basic research.

The understanding of charge movement in silica may seem to be a somewhat esoteric field; but, for obvious reasons, it is increasingly important to convey this knowledge to electronics undergraduates and to the growing numbers of workers entering the broader field of the development of "electronic systems on a chip". This book fills such a need. It is, for example, the first time I have seen a large

and carefully selected set of references on instability research.

Such research is carried out using a limited number of specialized electrical and optical techniques. These are outlined in 30 concise pages. The theoretical principles required include those of defect solid state physics, surface physics, lattice dynamics, conduction physics and molecular quantum mechanics, and treatment of these concepts is woven into the subsequent chapters. Among the research findings discussed are recent developments in surface states, mobile ions, polarization, hole trapping, electron trapping and electrical breakdown.

For many of the intended audience the book will be tough going. Readers are provided with reference to a few of the best, general solid-state technology textbooks, but the style of writing is somewhat flat and the graphics are unimaginative, coming, in the main,

directly from research papers or other textbooks. Nonetheless, the author does address all of the important conduction and trapping effects which determine whether or not a given MOS device will be "good".

The exploitation of conduction and trapping effects in the MOS system is only just beginning, and the practical aspects of instability research reinforce the usefulness of a review such as that provided by Dr Davis. The use of new and more subtle effects in MOS devices not only ensures "good" (i.e. stable) devices, but also leads to ones which are "better" than those of rival makers. In this context, it must not be forgotten that microelectronic devices are used not only in parlour games but also in machine tools, racing cars and missiles. The country which possesses the fastest and best microcircuits will perhaps win not only commercial success but also more explicitly gladiatorial contests. □

Andrew Holmes-Siedle is Consultant to the Physics Department of the Fulmer Research Institute, Stoke Poges, Buckinghamshire.

Statistical expertise on the shelf

P.T. Saunders

Biometry: The Principles and Practice of Statistics in Biological Research, 2nd Edn. By Robert R. Sokal and James F. Rohlf. Pp.859 plus tables. ISBN 0-7167-1254-7. (W.H. Freeman: 1981.) £19.50, \$31.95. Tables hbk £16.30, \$22; pbk £6.95, \$9.95.

FOR some years now, Sokal and Rohlf's *Biometry* has enjoyed considerable success as a statistics text for biologists. Knowing this, my first impressions of the second edition were less favourable than I had expected. While the opening chapters are written in an attractive enough style and deal with the right sort of material, I found them too long and diffuse. The authors seem so intent on ensuring that every aspect of the subject is covered that they leave the reader in danger of not seeing the wood for the trees. To give just one example, the crucial idea of the testing of hypotheses seems to get lost in a long section devoted mainly to Type I and Type II errors and power functions. It is not that the latter are not worth discussing, though they are relatively less important in a text in which derivations are omitted, but that the balance is wrong.

But while I have some reservations about the introductory material, the remainder of the book, which consists of detailed accounts of a large number of statistical techniques applicable to a wide variety of problems, is very good indeed. The authors make excellent use of boxes, set aside from the main text, in which they show precisely how each test is to be carried out. They

provide sample calculations for the benefit of those readers who are uncomfortable with subscripts and summation signs; more mathematically minded users will find them invaluable for testing computer programs. There are also informal discussions of the ideas behind the tests and other relevant points.

A number of improvements have been incorporated into the new edition and some additional material has been included, mainly concerned with correlation and regression. Inevitably the coverage is still not complete; for example, the authors have judged that their intended readership will be able to manage without such topics as sequential testing and Bayesian techniques. They have also chosen to publish the statistical tables separately instead of binding them with the book, which has its advantages so long as you can keep the two together.

Biometry would not be my first choice for someone with no previous knowledge of statistics, though it is certainly well worth considering. I would, however, strongly recommend it for anyone who already knows a bit about the subject and now finds himself in need of something beyond what he was taught. It is not quite as good as having a friendly statistician across the hall, but it can come surprisingly close. □

Peter Saunders is a Lecturer in Mathematics at Queen Elizabeth College, University of London.

● Another recent textbook on MOS devices is Oliver McCarthy's *MOS Device and Circuit Design*. The book, primarily aimed at students of electronic engineering, but also intended for practising engineers, is published by Wiley at £19.50, \$44.95.